COSMOLOGICAL INFORMATION FROM REDSHIFT SURVEYS

William Edmund Ballinger

Presented for the Degree of Doctor of Philosophy
The University of Edinburgh
1997

This thesis is dedicated to Mum, to Dad and to Catherine.

ACKNOWLEDGEMENTS

Many people have conspired to make the past four years of my life enjoyable and worthwhile. There now follows a somewhat gushy, but nevertheless honest and heartfelt list of thankyous.

Lots of thanks to Catherine for years of love and support. In particular, for the truly remarkable patience and tolerance on display while Grumpy here wrote up his thesis. This support and encouragement is particularly impressive from someone who, like most normal people of this Earth, tends to regard Ω_0 as just some Greek letter with a little zero stuck to the bottom.

Many thanks to Mum and Dad, for lots of love, support, and quite a bit of money. Also for taking me out for nice meals and taking me to the Opera (which I couldn't possibly afford), not to mention the icky childbirth bit. Sticking to immediate family, I must also thank my brother Geoff for letting me use his washing machine, live in his flat for a while, make fun of his hobbies, and just for being a friendly and supportive big brother in the best traditions of siblinghood.

I'd like to thank my supervisors, Alan Heavens, Andy Taylor and John Peacock for giving lots of support, advice and encouragement. I feel I have had a productive few years, with even a few publications under my belt, and none of this would have been possible without their assistance and prodding. Lots of thanks in particular to Alan for spending time going through my thesis drafts, and to Andy for bearing the brunt of my general cosmology queries and sometimes tangential scientific discussions. In addition to this trio, I'd like to thank my other collaborators, Helen Tadros and Will Saunders,

for being friendly and helpful and fun to work with.

I've enjoyed my (long) time at the Observatory, and the denizens have proved to be friendly. My fellow postgrad students have been a splendid lot. In particular, thanks to Mike, for being a great friend, telling me all the juicy gossip about his love life, and tolerating my cheap jibes about his school; Alison, for being a lovely wife and fellow cosmologist, and for showing what lies beneath her sweet and innocent exterior during the shopping incident in Southampton; Doddsy, for being a cuddly office mate with an occasional beard; Scary Rach, for being so huggable; Steve McNally, for sending me somewhat original emails; and Jo, for being bouncy and lively beyond the call of duty. Oh yeah, sorry for almost strangling Mat. Among the grown-ups, I'd particularly like to thank Suzie for giving me lots of lifts.

Further afield, I'd like to thank some other people for making life more interesting and occasionally decadent. They are (in no particular order): Gordon, Pete, Owen, Anjo, Catherine M, Kirstie, The White Family, Fleckie, ...this is like an Oscar speech..., Huddy, Catherine's Parents, their dogs, Al, Helen(s), Sara, Bondi and whoever invented aniseed balls. Also, the rest of my family – Grandparents, aunts, uncles *etc.* – particularly for remembering my birthday even though I usually haven't a clue when any of them have theirs.

I have no doubt that there are many other people whom I have forgotten to thank in this final rush to write-up. If you, the reader, feel you are one of these people, feel free to add your own personal paragraph in the space below.

This thesis is my own composition except where indicated in the text.

November 9, 1998

Contents

1 INTRODUCTION 1

1.1 Big Bang Cosmology 1

1.1.1 The Cosmological Principle 1

1.1.2 The Robertson Walker Metric 2

1.1.3 Expansion, Redshift and Thermodynamics 2

1.1.4 The Friedmann Equations 5

1.1.5 Horizons 8

1.2 The Hot Big Bang 8

1.2.1 A Brief History 9

1.2.2 Inflation 10

1.2.3 Nucleosynthesis 12

1.2.4 The Cosmic Microwave Background 13

1.3 Dark Matter 14

1.4 Structure Formation and Evolution 16

1.4.1 Correlation Function and Power Spectrum 17

1.4.2 Gaussian Random Fields 19

1.4.3 The Primordial Origin of Structure 20

1.4.4 The Pre-Recombination Era – The Transfer Function 21

1.4.5 Growth of Perturbations under Gravity 23

1.4.6 Bias 27

1.5 Measuring Ω_0 with Galaxy Surveys 28

1.5.1 Direct Peculiar Velocity Measurements and POTENT 29

1.5.2 The Dipole 30

1.6 Thesis Overview 31

2 REDSHIFT-SPACE DISTORTIONS **33**

2.1 Qualitative Effects of Redshift Distortions on the density field 35

2.2 Redshift Distortions in the Linear Regime 37

2.3 Non-Linear Redshift Distortions 41

2.3.1 Incoherent Velocities – The Fingers of God 41

2.4 Estimating β from Power-Spectrum Anisotropies 43

2.4.1 Linear theory power-spectrum multipoles 44

2.4.2 Fingers of God 45

2.4.3 Results 47

3 THE COSMOLOGICAL CONSTANT 53

3.1 Introduction 53

3.2 Cosmological Models and Clustering Anisotropy 55

3.3 Redshift Distortions 59

3.3.1 Linear modes 59

3.3.2 Fingers of God 60

3.3.3 The general case 61

3.3.4 Geometry from evolution of β 64

3.4 Maximum-Likelihood Fitting 65

3.4.1 Covariances and correlation matrix 67

3.4.2 Evolutionary effects 72

3.5 Applications 72

3.5.1 Galaxies 72

3.5.2 Quasars 73

3.6 Summary and Conclusions 75

4 SPHERICAL HARMONIC ANALYSIS – THEORY 78

4.1 Formalism and Transform 80

4.2 Redshift Surveys 81

4.2.1 Small–Scale Velocity Structure 85

4.2.2 Division into Radial and Angular Selection 88

4.3 Statistics 89

4.3.1 Covariance matrix terms – real and imaginary parts 89

4.3.2 Shot noise 91

4.3.3 Likelihood analysis 92

4.3.4 Optimal Weighting of the data 94

4.4 Discussion 97

5 SPHERICAL HARMONIC ANALYSIS – APPLICATIONS 98

5.1 Parameter Estimation Approach 99

5.1.1 Stepwise Maximum Likelihood 100

5.2 Tests on Simulated Surveys 102

5.2.1 Numerical simulations 102

5.2.2 Construction of mock PSCz surveys 104

5.2.3 Simulation Results – Fixed $P(k)$ 105

5.2.4 Simulation Results – Fitting $P(k)$ 110

5.3 Redshift Survey Results 111

5.3.1 1.2Jy Survey – Non-parametric Power Spectrum 111

5.3.2 PSCz Survey 115

5.4 Conclusions 124

6 DATA COMPRESSION 127

6.1 Maximum likelihood 128

6.1.1 Errors and Optimum Properties 129

6.1.2 Gaussian Data 130

6.2 Optimal Data Compression: Karhunen-Loève Methods 132

6.2.1 The Optimisation Problem 132

6.2.2 When the Mean is Known 134

6.2.3 Estimating several parameters at once 137

6.3 Correlated Parameters 138

6.3.1 The Problem 139

6.3.2 Parameter Eigenvector Optimisation 139

6.4 Tests on Simulations 143

6.5 Discussion 144

7 CONCLUSIONS 145

A Moments of the Power Spectrum Including Nonlinear Distortions 152

A.1 Moments for Gaussian Model 152

A.2 Moments of exponential model 154

B Fourier Transforms 156

C Selection functions and Weighting schemes 158

D Covariance matrices 160

D.1 Example 1: $\beta = 1$; $n = -1.5$; $R = 0$; $N = 1120$ 161

D.2 Example 2: $\beta = 0.3$; $n = -1.5$; $R = 0$; $N = 3210$ 162

D.3 Example 3: $\beta = 1$; $n = -1$; $R = 0$; $N = 715$ 163

D.4 Example 4: $\beta = 1$; $n = -1.5$; $R = 1000$; $N = 855\,000$ 164

E Window Functions 165

E.1 Practical calculation of window functions 165

E.2 Real and imaginary parts 167

F Orthogonality Relations 169

List of Figures

1.1 Transfer functions for example adiabatic models. The baryonic model does not scale exactly with Ωh^2, the case here is for $\Omega = 1$, $h = 0.5$. . . 22

2.1 The compression of a large scale overdensity along the line of sight in redshift space. The arrows show the coherent infall caused by the overdensity. 35

2.2 A 'Finger of God': the stretching of a virialised cluster along the line of sight in redshift space. The arrows denote the random motions of galaxies in the cluster. 36

2.3 Power spectrum contours for the purely linear formula (equation (2.5)) with $\beta = 0.5$. The power spectrum is boosted along the line of sight, stretching the contours out in the $k_{||}$ direction. The real space contours would be quarter circles due to isotropy. The assumed real-space power spectrum is a simple power law $P(k) \propto k^{-1.5}$. Note that $P(k, \mu)$ has its highest value in the bottom left corner. The contours are spaced in powers of two. 40

2.4 Contours of $P(k)$ with $\beta = 0.5$ including the effects of the 'Fingers of God' with $\sigma_p = 450\text{km s}^{-1}$. The thick contours are for the the Gaussian model (equation (2.21)), the thin contours for the exponential model (equation (2.22)). The power spectrum is suppressed along the line of sight at large k (small scale), squashing the contours out in the $k_{||}$ direction. At small k (bottom left corner) the linear boost effect dominates again. 47

2.5 The ratio of estimated to true β plotted against wavelength (scale), with $\beta = 0.5$ and 'Fingers of God' models with $\sigma_p = 450\text{km s}^{-1}$. The thick curve is for the the Gaussian model, the thin curve for the exponential model. 48

2.6 The β ratio curves from the CFW94 simulations with $\beta = 0.5$ and a measured $\sigma_p = 600\text{km s}^{-1}$. The dotted curves are for smoothed particle velocities and should be ignored here. 49

2.7 β ratio curve with $\beta = 0.5$ and $\sigma_p = 600\text{km s}^{-1}$. 50

2.8 CFW94 β ratio N-body results with $\beta = 0.24$ and $\sigma_p = 250\text{km s}^{-1}$. Ignore the dotted curves. 51

2.9 β ratio curve with $\beta = 0.24$ and $\sigma_p = 250\text{km s}^{-1}$. 52

2.10 β ratio curve with $\beta = 0.24$ and $\sigma_p = 160\text{km s}^{-1}$. 52

3.1 (a) Flattening factor F as a function of redshift for flat models ($\Omega_m + \Omega_\Lambda = 1$). The curves range from $\Omega_\Lambda = 1$ (top curve) to $\Omega_\Lambda = -1$ in steps of 0.1. (b) F against redshift for models with no cosmological constant. Ω_m varies from 0 (top) to 2, again in steps of 0.1. (c) F against redshift for models with a fixed mass density $\Omega_m = 0.2$ and a cosmological constant. Ω_Λ varies as for (a). 58

3.2 Contours of the power spectrum in the $k_{||}$ and $k_{\perp}$ plane, assuming a power-law index $n = -1.5$ for the true power spectrum. The contour interval is $\Delta \ln P = 1/2$. The models are linear (Kaiser) redshift distortion only with $\beta = 0.5$ (full contours) and geometric squashing effect only, with $F = 1.5$ (dotted contours). 60

3.3 Contours for both linear and non-linear redshift distortions with $\beta = 0.5$, and $\sigma_{\rm p} = 350$ km s^{-1} (again for $n = -1.5$). The solid contours are for exponential small scale velocity distribution, the dotted Gaussian. The contour interval is $\Delta \ln P = 1/2$. The stretching of contours of the correlation function along the line of sight leads to a reduction in line of sight power on small scales (large k) – see also Cole, Fisher & Weinberg (1994). 62

3.4 Solid contours for a model with $F = 1.1$, $\beta = 0.5$ and $\sigma_{\rm p} = 350$ km s^{-1}. Dotted contours are for the best fit to this model with redshift distortions only ($F = 1$). 63

3.5 Expected contours of likelihood in the $F-\beta$ plane, for the case of a survey with negligible shot noise, for three values of β. The scaling is set so that $\sigma(\beta) = 0.1$ for $\beta = 1$. At each (β, F) point, the amplitude distribution has been integrated over. The contour interval is $\Delta \ln \mathcal{L} = 1/2$. The position of the true values of the parameters is marked by the cross; contours are shown for $\beta = 0.3$, 0.6 and 1.0. 70

3.6 Expected contours of likelihood in the $F-\beta$ plane, as for Fig. 3.5 except now for the case of a survey dominated by shot noise. 71

3.7 Expected contours of likelihood in the $\Omega_\Lambda - \beta$ plane for a 2df-type galaxy redshift survey of 250 000 galaxies to $B_{\rm J} < 19.5$, obtained by combining cubic subsamples over a range of redshifts. Spatial flatness ($\Omega_{\rm m}+\Omega_\Lambda = 1$) is assumed. The models shown assume true values for β of 0.25 and 0.75, and $\Omega_\Lambda = 0$, as indicated by the crosses. 74

3.8 Expected contours of likelihood in the $\Omega_\Lambda - \beta$ plane for a 2df-type quasar survey of 30000 quasars with $z < 2.2$. Spatial flatness ($\Omega_m + \Omega_\Lambda = 1$) is assumed. The models shown assume true values for β of 0.25 and 0.75, and $\Omega_\Lambda = 0$, as indicated by the crosses. 76

5.1 Figure shows contours of likelihood in the β - $\Delta(k = 0.1)$ plane for 9 realizations of the PSCz redshift survey drawn from SCDM N-body simulations. The contours are plotted at intervals of $\delta \ln \mathcal{L} = 0.5$. The x axis is plotted in steps of 0.1 in $\Delta(k = 0.1)$. The vertical cross marks the true parameter values of $\Delta_{\rm true} = 0.54$, $\beta_{\rm true} = 1$. 106

5.2 Contours of likelihood in the β - $\Delta(k = 0.1)$ plane for 9 realizations of the PSCz redshift survey drawn from LCDM N-body simulations. Again, contours are plotted at intervals of $\delta \ln \mathcal{L} = 0.5$ and the x axis is plotted in steps of 0.1 in $\Delta(k = 0.1)$. The vertical cross marks the true parameter values of $\Delta_{\rm true} = 0.64$, $\beta_{\rm true} = 0.38$. 107

5.3 Contours of likelihood in the β - $\Delta(k = 0.1)$ plane for 9 realizations of the PSCz redshift survey drawn from MDM N-body simulations. The vertical cross marks the true parameter values of $\Delta_{\rm true} = 0.49$, $\beta_{\rm true} = 1$. 108

5.4 Figure shows the stepwise power spectrum estimation for the nine LCDM simulations with $l_{\rm max} = 19$, $n_{\rm max} = 7$, $r_{\rm max} = 300\,h^{-1}\,{\rm Mpc}$ and $k_{\rm max} = 0.073\,h\,{\rm Mpc}^{-1}$. The corresponding estimate of β is given alongside each curve. 109

5.5 The LCDM power spectrum estimates averaged together, showing good agreement with true power spectrum, and the correct $\beta_{LCDM} = 0.38$. . 110

5.6 The real-space power spectrum of *IRAS* 1.2-Jy galaxies. The solid line is the extrapolation of a Harrison-Zeldovich spectrum $P(k) \propto k$ from *COBE* measurements, assuming a quadrupole of $20\mu K$. The dotted line is mixed dark matter with 13% neutrinos (van Dalen & Schaefer 1992; Taylor & Rowan-Robinson 1992). The dashed line is the two-power law fit of Peacock & Dodds (1994) to a variety of catalogue data. All curves assume $\Omega_0 = 1$. 114

5.7 The real-space power spectrum of *IRAS* 1.2-Jy galaxies, as in Fig. 1, compared with redshift-space power spectra from the same survey (Fisher *et al.* 1993), and the *IRAS* QDOT survey, also in redshift-space (Feldman *et al.* 1994). The real-space power spectrum deconvolved from the APM catalogue (Baugh & Efstathiou 1994) is also shown. 116

5.8 The sky distribution of the PSCz sample used in this analysis. 117

5.9 Contours of likelihood for β and $\Delta_{0.1}$ for the PSCz survey. The boundary is at $r_{\max} = 240\,h^{-1}\,\mathrm{Mpc}$, with modes analysed up to $k_{\max} = 0.1\ h\,\mathrm{Mpc}^{-1}$, $l_{\max} = 21$ and $n_{\max} = 8$. The contours are separated by intervals of $\Delta \ln \mathcal{L} = 0.5$. 118

5.10 As for Fig 5.9, but with the addition of the Local-Group correction of equation (4.19). This correction has a negligible effect. 119

5.11 Contours of likelihood for β and $\Delta_{0.1}$ for the PSCz survey. The boundary is at $r_{\max} = 270\,h^{-1}\,\mathrm{Mpc}$, with modes analysed up to $k_{\max} = 0.89\ h\,\mathrm{Mpc}^{-1}$, $l_{\max} = 21$ and $n_{\max} = 8$. The contours are separated by intervals of $\Delta \ln \mathcal{L} = 0.5$. 120

5.12 Likelihood contours with $k_{\max} = 0.08\,h^{-1}\,\mathrm{Mpc}$ for (a) $r_{\max} = 240\,h^{-1}\,\mathrm{Mpc}$, (b) $r_{\max} = 270\,h^{-1}\,\mathrm{Mpc}$, (c) $r_{\max} = 300\,h^{-1}\,\mathrm{Mpc}$. 121

5.13 The real-space power spectrum of the PSCz redshift survey in dimensionless units. The curves are CDM models with $\Gamma = 0.2$ (dotted), 0.3 (dashed). 122

5.14 The real-space power spectrum of the PSCz (stars and cross-arrow) and 1.2Jy (triangles and arrow) redshift surveys. 123

6.1 An artificial (mean) likelihood contour for two highly correlated parameters. The conditional error on parameter θ_2 is given by the interval CC', the marginal error by MM'; see Press *et al.* (1992). The mean values are denoted by $\bar{\theta_1}$ and $\bar{\theta_2}$. 140

6.2 Illustration of the constraints imposed by different compression methods on the toy model of Fig 6.1. The dotted lines correspond to the conditional-error constraints imposed by the SVD method of TTH – clearly these have little influence on the length of the ellipse. The dashed line indicates the constraint imposed by the parameter-eigenvector method, a direct control of the length of the long axis itself. 141

6.3 Illustration of data compression with different algorithms. Top left: 'Full' dataset of 508 modes (for details of parameters etc, see text). Top right: Best 320 modes optimised for measuring β. Bottom left: Best 320 modes from SVD application to modes optimised for β and Δ. Bottom right: Best 320 modes for optimising along the likelihood ridge axis. Likelihood contours are separated by 0.5 in natural log. 143

List of Tables

5.1 Results of applying the two-parameter spherical-harmonic model to mock PSCz redshift surveys drawn from three cosmological models. Columns 1 and 3 give the recovered values of β and $\Delta(k = 0.1)$, averaged over nine realizations. These are to be compared with the true values of the parameters, given in columns 2 and 4. 110

5.2 Comparison of internal and external parameter errors from the 3 simulations. 111

5.3 Power spectrum measurements and limits for the 1.2Jy Survey. 115

5.4 Power spectrum measurements and limits for the PSCz Survey. 124

D.1 Three-parameter model. 161

D.2 Four-parameter model. 161

D.3 Three-parameter model. 162

D.4 Four-parameter model. 162

D.5 Three-parameter model. 163

D.6 Four-parameter model. 163

D.7 Three-parameter model. 164

D.8 Four-parameter model. 164

ABSTRACT

Structure observed in redshift surveys of galaxies contains information about fundamental properties of the universe, but these maps are distorted by the peculiar motions of galaxies. This distortion can, however, be accounted for and used to help constrain the density parameter Ω_0, usually via $\beta = \Omega^{0.6}/b$ where b is the bias parameter.

In chapter 1 a brief introduction is given to cosmology in general and to the origin, evolution and statistical analysis of large-scale structure in particular. Methods of measuring Ω_0 from large-scale velocities are briefly reviewed.

Chapter 2 introduces redshift distortions, the alteration of 3-dimensional redshift survey maps by peculiar velocities. Linear and nonlinear effects are reviewed both qualitatively and quantitatively. A simple nonlinear correction to a linear theory β estimator is tested and found to be moderately successful. It is clear, however, that for linearity to be reliable, very large scale modes must be analysed which subtend large angles in present surveys, preventing the use of conventional Fourier analysis. Instead spherical harmonic analysis must be used, which is pursued in chapters 4 and 5.

At high redshift the assumption of the incorrect cosmological model can lead to an additional geometric effect which can be confused with redshift distortions. If these can be disentangled, limits can be placed on the cosmological constant Λ in a way which is independent of source evolution. Chapter 3 introduces a detailed power-spectrum model including Λ with linear and nonlinear redshift distortions. The effects of evolution of the bias parameter are considered and a full statistical analysis is performed, showing that the next generation of redshift surveys may be just about capable of putting limits

on Λ.

A spherical harmonic and spherical Bessel function transform is introduced in chapter 4. Following and refining the analysis of Heavens & Taylor (1995), the effects of linear and nonlinear redshift distortions are modelled along with the effects of incomplete sky coverage and a radial selection function. The equivalence of this method to a conventional Fourier analysis and the advantages this entails are discussed. In chapter 5 the methods are applied to the *IRAS* 1.2Jy survey and the new PSCz survey. A non-parametric measurement of the shape and amplitude of the real-space (*i.e.* undistorted) power spectrum was introduced and applied to both surveys. In both spectra there is clear evidence for a turnover and Cold Dark Matter models fit fairly well, although there is marginal evidence for a tighter break. In addition, the values $\beta = 1.04 \pm 0.3$ and $\beta = 0.61 \pm 0.17$ (marginal errors) were measured respectively for the two surveys, simultaneously with the power spectrum. The latter result – from the superior survey – implies that *IRAS* galaxies must be biased if a flat $\Omega_0 = 1$ universe is required.

The likelihood methods discussed and/or used in chapters 3-5 can be computationally expensive to carry out, requiring the repeated inversion of large matrices. In the future, giant datasets could make the task of parameter estimation almost impossible. To deal with this, a method for compressing datasets while retaining information about *multiple*, *correlated* parameters is introduced and tested. The method appears to be very promising, and should prove very useful if applied to future surveys.

Chapter 1

INTRODUCTION

This thesis is concerned with the measurement of fundamental cosmological information from redshift surveys of galaxies. This chapter includes a brief review of cosmology, and then a more detailed discussion of the measurement and evolution of large-scale structure, and a review of methods of constraining the density parameter Ω_0

1.1 Big Bang Cosmology

1.1.1 The Cosmological Principle

The Cosmological Principle states that the Universe is homogeneous and isotropic, which is equivalent to saying that it is isotropic everywhere. This clearly does not apply in detail: matter in the universe clusters into overdense and underdense regions. On larger and larger scales however, the density distribution starts to look smoother, although this is a subtle point – fluctuations in the *potential* are believed to be almost constant as a function of scale at the largest scales – see section (1.4.3). Isotropy is very well established by the evenness of radio source counts and the smoothness of the microwave background radiation. It would seem perverse to many people to assume that we are in some special place in the universe, so if it is isotropic here it is assumed to be isotropic everywhere which implies homogeneity.

1.1.2 The Robertson Walker Metric

Using the field equations of general relativity the spatial symmetry imposed by the cosmological principle leads to a very simple space-time metric known as the Robertson-Walker (RW) metric,

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - R^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right], \tag{1.1}$$

see Weinberg (1972). The high degree of symmetry also allows a simpler derivation – see Longair (1984). The time coordinate is *cosmic time*, as measured on the clocks of *fundamental observers*. These are free-falling observers at rest relative to their local region of the universe who observe the universe to be isotropic. The clocks can be synchronised by comparing the observed mean density of the universe in the two locations. The spatial part is the generalisation of the metric of a sphere, with an expansion factor $R(t)$, which has the dimensions of length – r is dimensionless. The metric is often expressed with a dimensionless scalefactor $a(t) = R(t)/R(t_0)$, set to unity at the present epoch. The coordinates can be scaled so that the curvature constant k can only have values -1,0 or 1; corresponding to an open, flat or closed spatial geometry. If the metric is recast using the dimensionless scalefactor (note that the r here is different),

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\left[\frac{\mathrm{d}r^2}{1-Kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right], \tag{1.2}$$

K is no longer simple and dimensionless, but the new r is the *comoving angular diameter distance*.

1.1.3 Expansion, Redshift and Thermodynamics

One of the most important cosmological discoveries made this century is the cosmological redshift – photons from distant galaxies have their wavelengths stretched. The redshifts, defined as

$$z \equiv \frac{\Delta\lambda}{\lambda}, \tag{1.3}$$

can be interpreted (locally) as being caused by a recession velocity ($z \propto v$), and were found to be proportional to their distances from the observer (Hubble 1929) *i.e.* they

are caused by a uniform expansion. The expansion rate is usually defined in terms of the Hubble parameter,

$$H \equiv \frac{\dot{a}}{a}, \tag{1.4}$$

which gives the Hubble law relating distance (not comoving) to recession velocity

$$v_{\rm r} = Hd. \tag{1.5}$$

Note that strictly speaking d distance is the *proper distance* defined by measurements made with a chain of rulers held by fundamental observers at fixed t. The local value of the Hubble parameter (denoted H_0) is one of the fundamental parameters of cosmology, and is the subject of ongoing debate – see *e.g.* Rowan-Robinson (1985). To allow for this uncertainty a fudge factor is frequently introduced into formulae,

$$h \equiv \frac{H_0}{100}, \tag{1.6}$$

where H_0 is in units of km s^{-1}Mpc^{-1}.

At high redshift recession velocity is not well defined and the redshift effect must be calculated more carefully. If a wavecrest of a light wave is emitted at position $(r_{\rm e}, 0, 0)$ at time $t_{\rm e}$ and observed at the origin at time $t_{\rm o}$, it follows from the metric equation (1.2) that

$$\int_0^{r_{\rm e}} \frac{{\rm d}r}{\sqrt{1-Kr^2}} = \int_{t_{\rm e}}^{t_{\rm o}} \frac{{\rm d}t}{a(t)}. \tag{1.7}$$

The next crest is emitted at $t_{\rm e} + \Delta t_{\rm e}$ and observed at $t_{\rm o} + \Delta t_{\rm o}$, then

$$\int_0^{r_{\rm e}} \frac{{\rm d}r}{\sqrt{1-Kr^2}} = \int_{t_{\rm e}+\Delta t_{\rm e}}^{t_{\rm o}+\Delta t_{\rm o}} \frac{{\rm d}t}{a(t)}. \tag{1.8}$$

Assuming the expansion rate is slow relative to the frequency of the light, $a(t) \simeq a(t+\Delta t)$, it follows that

$$\frac{\Delta t_{\rm e}}{a(t_{\rm e})} = \frac{\Delta t_{\rm o}}{a(t_{\rm o})}, \tag{1.9}$$

i.e. the interval between crests increases linearly with scalefactor, and hence so does the wavelength. The redshift is defined as

$$z \equiv \frac{\lambda_{\rm o} - \lambda_{\rm e}}{\lambda_{\rm e}} = \frac{a(t_{\rm o})}{a(t_{\rm e})} - 1, \tag{1.10}$$

so it is a measure of the scalefactor at the time of emission:

$$a(t_{\rm e}) = \frac{1}{1+z}. \tag{1.11}$$

Modelling the content of the universe as a uniform gas, we have

$$\mathrm{d}(\rho c^2 a^3) = -p\mathrm{d}a^3, \tag{1.12}$$

where ρ is density of material (matter, radiation *etc.*) and p is the pressure; this is simply the first law of thermodynamics for an adiabatic system: $\mathrm{d}U = -p\mathrm{d}V$. The universe is clearly adiabatic because of homogeneity – statistically, any region has as much heat flowing in as flowing out. Equation (1.12) can be manipulated to give

$$c^2 \frac{\mathrm{d}\rho}{\mathrm{d}a} + 3\left(\frac{p + \rho c^2}{a}\right) = 0. \tag{1.13}$$

The equation of state of a non-relativistic ideal gas is given by

$$p_\mathrm{g} = \frac{3}{5}\rho_\mathrm{g} v_\mathrm{s}^2 \tag{1.14}$$

where v_s is the sound speed. Substituting into equation (1.13) and noting that $v_\mathrm{s} \ll c$ gives

$$\rho_\mathrm{g} \propto a^{-3}. \tag{1.15}$$

This is easily understood, ρ_g is just the rest mass density, and so decreases by the same factor by which the volume increases. For radiation (and relativistic particles),

$$p_\mathrm{r} = \frac{1}{3}\rho_\mathrm{r} c^2, \tag{1.16}$$

giving

$$\rho_\mathrm{r} \propto a^{-4}, \tag{1.17}$$

the extra factor of a^{-1} coming from the redshift. It is clear from this that the relative energy density of photons to non-relativistic matter falls as the universe expands, so that the present universe is matter dominated, but at very high redshifts it is radiation dominated. The microwave background radiation appears to obey a black body spectrum to a high degree of accuracy – see section 1.2.4. The energy density per unit frequency of black-body radiation with temperature T_0 at the present epoch ($a = 1$) is given by the Planck function

$$u_0(\nu_0)\mathrm{d}\nu_0 = \frac{8\pi h}{c^3} \frac{\nu_0^3 \mathrm{d}\nu_0}{\exp\left(\frac{h\nu_0}{kT_0}\right) - 1}. \tag{1.18}$$

Photon numbers are preserved in a comoving volume, they are merely redshifted. Hence mapping the frequency back and applying the redshift relations $\nu_0 \rightarrow \nu = \nu_0(1+z)$ and $\mathrm{d}\nu_0 \rightarrow \mathrm{d}\nu = \mathrm{d}\nu_0(1+z)$ the energy density is given by

$$u(\nu)\mathrm{d}\nu = \frac{8\pi h}{c^3} \frac{\nu^3 \mathrm{d}\nu}{\exp\left(\frac{h\nu}{kT}\right) - 1}, \tag{1.19}$$

a black body with temperature $T = T_0(1+z)$. The black body form is preserved, so the microwave background could have been thermalised at high redshift.

If there is a vacuum energy, corresponding to a cosmological constant, then the energy density is constant:

$$\rho_\mathrm{v} = \text{constant} \tag{1.20}$$

which corresponds to a negative pressure,

$$p_\mathrm{v} = -\rho_\mathrm{v} c^2. \tag{1.21}$$

Notice that the relative dominance of vacuum energy increases with scalefactor, with

$$\frac{\rho_\mathrm{v}}{\rho_\mathrm{g}} \propto a^3. \tag{1.22}$$

1.1.4 The Friedmann Equations

The evolution of the scalefactor $a(t)$ can be derived from the GR field equations, but for a simple dust-filled universe only Newtonian Mechanics is required, see *e.g.* Longair (1984). The general result is this Newtonian dust result modified by an extra term $3p/c^2$. The evolution is described by the Friedmann equations,

$$\frac{\ddot{a}}{a} = -\frac{4\pi}{3}\left(\rho + 3\frac{p}{c^2}\right) \tag{1.23}$$

and

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{Kc^2}{a^2} = \frac{8\pi}{3} G\rho, \tag{1.24}$$

where the constant K is the same as appears in equation (1.2). These are not independent, the second can be derived from the first by integration if the adiabatic work done by expansion is considered – see equation (1.12). Equation (1.24) shows that for a spatially flat universe with $K = k = 0$

$$\rho = \frac{3H^2}{8\pi G} \equiv \rho_\mathrm{c}, \tag{1.25}$$

the critical density, remembering the definition of the Hubble parameter – $H = \dot{a}/a$. The ratio of the density to this critical density,

$$\Omega \equiv \frac{\rho}{\rho_c}, \tag{1.26}$$

is one of the fundamental parameters of cosmology. A universe with $\Omega = 1$ is called a critical universe, and is exactly balanced between two other cases. A closed universe has $\Omega > 1$ and $k = 1$, is spatially finite (the space slices have geometry corresponding to a sphere) and, in the absence of a cosmological constant, will expand, stop and recollapse in a finite time. If $\Omega < 1$ the universe is open, spatially infinite and will expand forever.

As discussed in section 1.1.3, the value of p as a function of ρ depends on the equation of state of the contents of the universe. The cosmological constant can be included in the Friedmann equations as an extra curvature term, as originally introduced by Einstein. This complexity is unnecessary; instead Λ can be regarded as a vacuum source term with $p = -\rho c^2$, giving a constant density independent of scalefactor $a(t)$. A universe dominated by a cosmological constant $\Omega_\Lambda = 1$ is called de Sitter space and equation (1.24) shows that it expands exponentially,

$$a(t) = \exp(Ht). \tag{1.27}$$

Assuming that the total density can be expressed as a sum of vacuum, matter (dust) and radiation contributions, the total variation with scalefactor is given as

$$\rho = \rho_m a^{-3} + \rho_r a^{-4} + \rho_v. \tag{1.28}$$

Ultrarelativistic particles can be treated as additional radiation density. It is clear by inspection that for large values of a, vacuum energy, if present, will tend to dominate the expansion. If there is no vacuum energy (as many workers believe) then the earlier phases are radiation dominated and later phases are matter dominated.

The curvature parameter K is fixed in an RW model, so a critical $\Omega = 1$ model is always critical, regardless of the equation of state. For other models, the evolution of Ω is more complicated. If the present density parameter is made up of radiation, matter

and vacuum contributions, $\Omega_0 = \Omega_r + \Omega_m + \Omega_\Lambda$, then substituting equation (1.28) into equation (1.24) and using $\Omega - 1 = Kc^2/(H^2a^2)$ gives

$$\Omega(a) - 1 = \frac{\Omega_0 - 1}{\Omega_\Lambda a^2 + \Omega_m a^{-1} + \Omega_r a^{-2} - \Omega_0 + 1}. \tag{1.29}$$

Again there is a fundamental difference between the effect of vacuum energy and of radiation or matter. If there is no vacuum energy, then $\Omega = 1$ is unstable and the universe will evolve away from this point, becoming either more open or more closed. Looking back to high redshift (low a) the universe will look more and more critical. Present estimates of the density parameter are in the range $0.2 < \Omega_0 \leq 1$, remarkably close to flat. This means that at the Planck time $t_P \simeq 10^{-43}$s – before which classical GR is expected to break down – $|\Omega - 1| \leq 10^{-55}$. This would seem to require an incredible degree of fine tuning in the initial conditions, and is known as 'the flatness problem' of classical cosmology.

Equation (1.29) shows that vacuum energy has the opposite effect – it drives Ω towards unity. In fact inflationary theory can solve the flatness problem by creating a temporary vacuum energy which flattens spacetime then disappears; this will be discussed in more detail in section 1.2.2. Models in which $\Lambda \neq 0$ at the present epoch will be discussed in more detail in chapter 3.

Even without inflationary models, many workers assumed that $\Omega_0 = 1$, which is known as the Einstein-de Sitter model if it is matter dominated. It seems the most natural and elegant solution, and is also the most convenient to work with, the space-slices are simply euclidean. This kind of appealing reasoning may seem to sidestep the fine-tuning argument given above; slightly open models, favoured by many observations, cannot fall back on such aesthetic arguments.

A flat model with non-zero vacuum energy of order $\Omega_\Lambda = 0.8$, as suggested by Efstathiou, Sutherland & Maddox (1990), fits in well with inflation but still requires some fine tuning. If this is the case, the universe is on the fringes of de Sitter space, with vacuum energy about to totally dominate the dynamics of the expansion – see equation (1.22) and equation (1.29). This model requires fine-tuning of a tiny vacuum energy at high redshift. It would again seem unlikely that observers would exist at a special epoch where matter and vacuum contributions are similar. A possible solution

based on the Anthropic Principle, the point that observers could only see the subset of parameter space which allows observers to exist, was suggested by Efstathiou (1995). The point is made that if vacuum dominated much earlier, structure would be unable to form due to the accelerated expansion, and so no observers would exist. This does not, however, explain why the vacuum energy is not undetectably small in the model.

1.1.5 Horizons

Causal processes in the universe can only occur between regions which can communicate with light signals. The Robertson-Walker metric can have a particle horizon, from where an observer can only receive signals emitted at the origin of time $t = 0$, and hence has no causal connection with any more distant regions. In the highly symmetric RW universe, particle horizon is at proper distance

$$R_{\rm PH}(t) = a(t) \int_0^t \frac{c{\rm d}t'}{a(t')}. \tag{1.30}$$

This integral does not converge for de Sitter space, so there is no particle horizon in this model (in fact there is no origin of time when $a(t_{\rm origin}) \equiv a(0) = 0$).

Another concept is the Hubble radius, loosely known as "the horizon" (this can lead to confusion). This is simply defined as

$$R_{\rm H} = \frac{c}{H}. \tag{1.31}$$

This is a measure of the effective distance across which causal processes can act at a given time. For cosmological models which had no inflationary phase this is roughly the same as the particle horizon; however, inflation can cause the particle horizon to be very much bigger than the Hubble radius – see section (1.2.2).

1.2 The Hot Big Bang

With the exception of some vacuum energy models, equation (1.24) implies that the universe has come from a state of infinite density ($a = 0$) at a finite time in the past.

However, GR is expected to break down at the Planck Time, $t_{\rm p} \simeq 10^{-43}s$ when quantum gravity effects take over. A "Hot" Big Bang model can be roughly defined as one in which the universe passed through a high temperature phase after $t_{\rm p}$ – hot enough for the standard primordial nucleosynthesis model to apply ($T \sim 10^{12}$K).

This introduction is quite brief and concentrates particularly on theory relevant to the density parameter Ω_0. See *e.g.* Kolb & Turner (1990) or Coles & Lucchin (1995) for further details.

1.2.1 A Brief History

The hot big bang model passes through a number of phases during expansion as the contents cool due to redshift. Note that for relativistic species $T \propto a^{-1}$ (unless the number of available particle states changes). The phases and events are as follows:

1. The Planck epoch, $t < 10^{-43}s$. The details of this time are particularly uncertain as there is no full theory of quantum gravity.

2. The GUT era. Up until $t \simeq 10^{-37}s$, $T \simeq 10^{15}$GeV electromagnetism, the weak nuclear force and the strong nuclear force are believed to be combined as one force in a *Grand Unified Theory*. At the end of this era there is a phase transition and the strong nuclear force becomes separate.

3. Inflation. Many workers believe a period of exponential expansion occurred around the time of the GUT phase transition. This will be dealt with in the next section.

4. Electroweak Era. The weak and electromagnetic forces are combined as one force until $t \simeq 10^{-11}s$.

5. The Quark Era. According to the Standard model hadrons are made up of quarks which are unbound at very high energies. At $t \simeq 10^{6}s$ these recombine in the *Quark-Hadron phase transition*, forming mesons (pions – quark-antiquark pairs) and baryons.

6. The Hadron Phase. This is a brief period before the pions annihilate when the temperature drops to $T \simeq 10^{12}$K.

7. The Lepton Phase. Lasting until electron-positron pairs annihilate at $T \simeq 5 \times 10^9$K, this period sees the freeze-out (decoupling) of neutrinos. Since the annihilation raises the temperature of the still-coupled components, the neutrinos remain at a ($\sim 40\%$) lower temperature.

8. Nucleosynthesis. Light nuclei are formed from protons and neutrons – this is dealt with in section 1.2.3

9. The Plasma phase. The primordial ions are tightly coupled to the radiation field by Thomson scattering. This affects the growth of baryon-dominated fluctuations – see section 1.4.4.

10. Matter-Radiation equality. The dynamics of the expansion change as matter becomes the dominant component (because $\rho_{\rm m}/\rho_{\rm r} \propto a(t)$). Sub-horizon structure growth begins if there is cold dark matter.

11. Recombination[1]. The temperature drops far enough that the ions and electrons combine to form neutral atoms. The cross section for photon absorption drops sharply and the two species decouple. The newly decoupled photons continue to redshift and become the microwave background, preserving information about the density fluctuations at decoupling – see section 1.2.4.

1.2.2 Inflation

There are a number of problems with the classical Hot Big Bang model:

1. *The Flatness Problem* alluded to earlier, where the observed local density is remarkably close to critical when this is an unstable equilibrium.

2. *The Cosmological Horizon Problem* – the microwave background has a remarkably uniform temperature across regions greater than the apparent particle horizon,

[1]This is, of course, a misnomer; the atoms were never combined in the first place. Typical.

i.e. regions which, in the standard matter dominated model, have never been in causal contact.

3. *The Origin of Structure Problem* – the observed structure in the universe is generally believed to have evolved from primordial fluctuations by gravitational instability. What is the origin of these fluctuations? How can these fluctuations be created on scales greater than the Horizon?

The third problem will be briefly addressed in section 1.4.3; the first two (and the second part of the third) can be explained simply in terms of a 'temporary cosmological constant' where the universe undergoes a phase of exponential expansion – the (comoving) Hubble length shrinks. This drives the universe to $\Omega = 1$ and greatly increases the size of the particle horizon – hence causal processes can have created the superhorizon homogeneity (problem 2 above) and inhomogeneity (problem 3) that is observed.

A number of inflationary models exist, some of which are built around the phase transitions in the early universe (*e.g.* at the end of the GUT era). In an analogy with phenomena such as supercooled steam, the universe is trapped in a false vacuum which gives rise to a vacuum energy field – the inflaton – which drives inflation. In most scenarios this is very close to de Sitter space, although ρ_v does evolve. There are some problem associated with these phase transitions, involving fine-tuning arguments. An alternative suggested by Linde (1983) is chaotic inflation, where the local universe simply starts with an energy field far from equilibrium which then evolves during the inflation period – no phase transition involved. At the end of the inflation phase, the energy of the inflaton transferred to particle fields which are coupled to it.

Most inflationary models predict $\Omega = 1$ – to avoid this the period of inflation would have to end at a very precise time and again fine tuning would be required. There are some more contrived models with a second inflationary period which allow an open universe with rather less fine tuning – see *e.g.* Bucher, Goldhaber & Turok (1995).

1.2.3 Nucleosynthesis

A constraint on the density of Baryons in the universe can be derived from arguments about nucleosynthesis in the early universe – see Copi, Schramm & Turner (1995) for a recent review. Only light nuclei, from deuterons to ^{7}Li, are produced from primordial protons and neutrons in this phase – the resonant triple alpha reaction which produces carbon from triplets of helium atoms requires far higher densities to proceed rapidly and can only take place efficiently in stellar interiors.

In primordial nucleosynthesis deuterium nuclei are formed from protons and neutrons,

$$n + p \rightleftharpoons d + \gamma, \tag{1.32}$$

which then combine to form helium

$$\begin{aligned} d + d &\rightarrow {}^3He + n, \\ {}^3He + d &\rightarrow {}^4He + p. \end{aligned}$$

The details of the reaction rates as a function of density are quite involved, but can be robustly calculated using numerical methods, as originally performed by Wagoner, Fowler & Hoyle (1967). The abundance of ^4He by mass is measured to be $Y = 0.22$-0.23 (*e.g.* Pagel *et al.* 1992), but this is rather insensitive to total baryon content – see *e.g.* Kolb & Turner (1990). The abundance of primordial deuterium, ^3He and ^7Li are more sensitive to the baryon density, but these nuclei are much more difficult to detect. There are problems detecting the mean density of these primordial species as their abundance varies between different regions and there is contamination from stellar nucleosynthesis. Nevertheless, the results yield a strong constraint on the abundance of baryons – Walker *et al.* (1991) found a very tight constraint using abundance measurements from a number of sources,

$$0.01 \leq \Omega_{\rm b} h^2 \leq 0.015. \tag{1.33}$$

More recent results, using observations of the d/H ratio from absorption lines (in the spectra of high redshift quasars) caused by high redshift gas clouds have found a slightly higher value,

$$0.019 \leq \Omega_{\rm b} h^2 \leq 0.030, \tag{1.34}$$

see Tytler, Fan & Burles (1996). This result is arguably closer to the primordial value – the high redshift gas is in clouds which do not appear to have formed stars, so the elements have probably not been processed since the nucleosynthesis phase. Even allowing for significant systematic effects this larger value is still far below most estimates of the total matter contribution $\Omega_{\rm m}$ – see section 1.3. Clearly the bulk of the matter in the universe must be non-baryonic.

1.2.4 The Cosmic Microwave Background

The primordial radiation from the time of decoupling at $z \sim 300$ was predicted by Gamow (1948) and detected serendipitously by Penzias & Wilson (1965). It is the relic radiation from the big bang, decoupled after recombination and left to adiabatically cool with the expansion.

Observations with the COBE satellite (Smoot *et al.* 1992) show that this radiation follows a Planck black body form (equation (1.19)) with extreme precision, and has a best-fit temperature of (Mather *et al.* 1994)

$$T = 2.726 \pm 0.010\mathrm{K}. \tag{1.35}$$

In addition to accurately pinning down the temperature and spectrum of the radiation, COBE was also the first instrument to unambiguously detect fluctuations in the radiation across the sky (previous experiments may have been detecting galactic emission). These correspond to density fluctuations at decoupling, and can be used to place limits on the normalisation and slope of the fluctuation spectrum. Future experiments (*e.g.* the proposed Planck Surveyor) will analyse the spectrum in more detail and place firm limits on many cosmological parameters, see White, Silk & Scott (1994) for a review. The actual relation between the observed black-body temperature fluctuations and mass overdensities involves several processes on different scales – see also Efstathiou (1990) for more details of the calculations. Large-scale fluctuations (superhorizon) are dominated by the Sachs-Wolfe effect (Sachs & Wolfe 1967), potential fluctuations cause photons to undergo a gravitational redshift – deeper potential wells redshift photons more. On smaller scales other effects dominate, doppler shifts caused by moving plasma

and adiabatic variations in temperature. An adiabatic overdensity causes the plasma to recombine at a later cosmological time, and so the photons have redshifted less by cosmological expansion by the time they are observed; similarly underdensities increase the redshift.

1.3 Dark Matter

A fundamental problem in astronomy is that matter can only be directly observed if it is emitting radiation for observers to detect. There is a very substantial body of indirect evidence for *dark matter*, material in addition to the photon-emitting stars, gas and dust – see *e.g.* Trimble (1987) and Binney & Tremaine (1987) for a review.

Dark matter was originally detected in the Coma cluster by Zwicky (1933). The high velocity dispersion implied a mass far greater than the apparent sum of its galaxy masses. More recently, hot X-ray emitting gas has been detected in clusters, but even this extra mass cannot account for the observed dynamics. Galaxies themselves show evidence for a halo of dark matter; in spiral galaxy discs the rotation curves (plots of rotational velocity against radius) are often flat for a considerable extent beyond the limits of the visible matter. The Kepler prediction is $v(r) \propto r^{-1/2}$, so there is clearly unseen matter surrounding the galaxies.

The relation between light and mass is usually quantified as a mass-to-light ratio; this is quoted in solar units so $(M/L)_\odot \equiv 1$. For a critical universe the mean value has to be very large,

$$\left.\frac{M}{L}\right|_{\rm crit} = 1500h^{+700}_{-400}, \tag{1.36}$$

see Efstathiou, Ellis & Peterson (1988); most direct estimates are high ($\gg 1$) but fall short of this. Dynamical measures of clusters consistently give a ratio of

$$\left.\frac{M}{L}\right|_{\rm cluster} \simeq 300h, \tag{1.37}$$

see Faber & Gallagher (1979) and Carlberg *et al.* (1996). At face value this implies a universe which is open, $\Omega \sim 0.2$, but galaxy clusters may not be representative of the universal M/L ratio – see the discussion of bias in section 1.4.6.

Gravitating matter can be detected directly by gravitational lensing – see *e.g.* Narayan & Bartelmann (1996). In the case of clusters, surface density profiles can be constrained from observations of giant arcs, weak shearing of images (Kaiser & Squires 1993) or number counts of background galaxies (Broadhurst, Taylor & Peacock 1995). See Mellier *et al.* (1996) for a review, including lists of results for giant arc and weak shear methods. The results tend to produce $M/L \sim 300h$, but there is a significant scatter. Additionally, if dark matter is composed of compact bodies it may be detectable from a microlensing effect – weak amplification of background sources – see the discussion of MACHOs below.

In all this work on cluster masses there may be a bias towards systems with low mass-to-light ratios because the clusters are often selected optically. There have been recent reports of "dark clusters", with few visible galaxies, so the mean M/L may be higher, exacerbating the problems for baryonic dark matter, see *e.g.* Jones *et al.* (1997).

The nature of the dark matter is a major unanswered question in cosmology. The large discrepancy between the value of $\Omega \sim 0.2$ for clusters and $\Omega_{\rm baryon} \sim 0.02$ from nucleosynthesis implies that the bulk of the dark matter is nonbaryonic, but see also the review by Carr (1994).

Dark matter can be divided into two broad classes:

1. Hot dark matter (HDM). These particles have relativistic velocities when they decouple, and their high velocities have a strong damping effect during structure formation and are difficult to reconcile with observations – see *e.g.* White, Frenk & Davis (1983) and White, Davis & Frenk (1984). A classic example of HDM is a massive neutrino.

2. Cold Dark Matter (CDM). These are non-relativistic at decoupling, and have a different effect on structure formation. No non-baryonic form of CDM is as well motivated by particle physics as the HDM neutrino as none has ever been detected, let alone had its mass measured. An alternative is to have dark matter in the form of MACHOs, Massive, Compact Halo Objects. These could be brown dwarfs – stars that are too small to burn hydrogen and hence are very cool and

hard to detect. Several groups have claimed detections of microlensing events in our galaxy, see *e.g.* Alcock *et al.* (1997). Like most MACHO candidates, these are baryonic and so there may be difficulties with the nucleosynthesis arguments of section (1.2.3). Another possible MACHO candidate is primordial black holes, which may be responsible for the long-term variation in quasar light curves via microlensing – see Hawkins (1993). If these are formed in the early universe before nucleosynthesis they avoid the low-$\Omega_{\rm baryon}$ constraint.

There are a number of direct and indirect methods to establish the nature of the unseen mass. Indirectly, the effects on structure formation can be analysed by measuring the present day clustering of galaxies – see section (1.4.4)

It is clear that very large scales must be probed in order to measure the density of a representative part of the universe. A large part of this thesis involves measuring the density parameter at very large scales, as well as the details of the clustering to try to discover the nature of the dark matter.

1.4 Structure Formation and Evolution

The analysis in previous sections has largely assumed that the matter in the universe is smoothly distributed; this is clearly untrue in detail. On many scales matter forms into clumps – galaxies, clusters of galaxies, stars, armchairs, gerbils etc. Analysis of large scale structure can provide information about cosmological parameters and the constituents of the universe.

The creation of the observed large-scale density field can be divided roughly into three sections; the primordial origin (from inflation or defects), the pre-recombination evolution, and the gravity-dominated post-recombination evolution.

Models of the origin of structure do not predict the exact details of the density field, but rather its statistical properties. To assess these models we need to find statistical measures of the clustering of objects.

See Peebles (1980), Peebles (1993) and Efstathiou (1990) for more details about large-scale structure.

1.4.1 Correlation Function and Power Spectrum

In analysis of large-scale structure, we are considering deviations from the mean density of the universe $\bar{\rho}$, which can be quantified as a fractional overdensity

$$\delta(\mathbf{x}) \equiv \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}}, \tag{1.38}$$

defining an 'overdensity' field in the universe. This field can be Fourier transformed

$$\delta_{\mathbf{k}} \equiv \int \delta(\mathbf{x}) \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}} \mathrm{d}^3 x \tag{1.39}$$

with inverse

$$\delta(\mathbf{x}) \equiv \frac{1}{(2\pi)^3} \int \delta_{\mathbf{k}} \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{x}} \mathrm{d}^3 k. \tag{1.40}$$

Calculating the evolution of the density field turns out to be easier in Fourier space, as the modes grow independently if the overdensities are small. Clearly this is isn't the case in real space, where overdensities are influenced by their local environment.

Many workers prefer to use a discrete transform, with the universe assumed to be periodic in some large cubic volume V, *e.g.*

$$\delta(\mathbf{x}) = \sum_{\mathbf{k}} \delta_{\mathbf{k}} \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{x}}. \tag{1.41}$$

See Bertschinger (1992) for details about conventions.

A common statistic for the density field is the correlation function (strictly the autocorrelation function), a function only of distance $x' = |x'|$ because of isotropy and independent of position $\mathbf{x}$ because of statistical homogeneity:

$$\xi(x') \equiv \langle \delta(\mathbf{x}) \delta(\mathbf{x} + \mathbf{x}') \rangle = \int \langle \delta_{\mathbf{k}} \delta_{\mathbf{k}'} \rangle \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}} \mathrm{e}^{i\mathbf{k}'\cdot(\mathbf{x}+\mathbf{x}')} \mathrm{d}^3 k \mathrm{d}^3 k', \tag{1.42}$$

where $\langle ... \rangle$ denotes ensemble averaging over different realisations of the density field. Strictly speaking, this corresponds to averaging over different universes; however, the density field is assumed to have the property of *ergodicity*, which means that spatial

averaging is equivalent to this (a simple non-ergodic field is a completely uniform one where the mean density is a random variable). In particular, a Gaussian random field is ergodic, see Adler (1981).

More intuitively, the correlation function can also be regarded as the fractional excess probability density of finding a galaxy a distance $\mathbf{x}'$ from another galaxy,

$$\delta P(\text{galaxy at } \mathbf{x}+\mathbf{x}'|\text{galaxy at } \mathbf{x}) = n^2[1+\xi(x')], \tag{1.43}$$

where n is the mean density.

The correlation of modes in Fourier space – covariance matrix if the transform is discrete – is a simpler object than the spatial equivalent,

$$\langle \delta_{\mathbf{k}} \delta_{\mathbf{k}'} \rangle \equiv (2\pi)^3 P(k) \delta^{\mathrm{D}}(\mathbf{k}+\mathbf{k}'). \tag{1.44}$$

Modes at different $\mathbf{k}$ are independent due to homogeneity – see Peebles (1980).

The reality of the overdensity field means that the Fourier coefficients satisfy the relation $\delta_{-\mathbf{k}} = \delta_{\mathbf{k}}^*$ and so equation (1.44) now becomes

$$\langle \delta_{\mathbf{k}} \delta_{\mathbf{k}'}^* \rangle \equiv (2\pi)^3 P(k) \delta^{\mathrm{D}}(\mathbf{k}-\mathbf{k}'). \tag{1.45}$$

The quantity $P(k)$, the variance on the k-space field, is known as the power spectrum. Isotropy and homogeneity demand that it depends only on the magnitude of $\mathbf{k}$.

The correlation function and the power spectrum form a Fourier pair,

$$\xi(x') = \frac{1}{(2\pi)^3} \int P(k) \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{x}'} \mathrm{d}^3k \tag{1.46}$$

and

$$P(k) = \int \xi(x') \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}'} \mathrm{d}^3x', \tag{1.47}$$

so in principle full knowledge of either of these functions contains complete information about the field

Another important quantity of interest is the *rms* mass fluctuation or variance of the spatial density field as it is sometime known:

$$\langle \delta^2(\mathbf{x}) \rangle \;=\; \xi(0) = \frac{1}{(2\pi)^3} \int P(k) \mathrm{d}^3k$$

$$= \int \frac{2}{\pi^2} k^3 P(k) \frac{\mathrm{d}k}{k}$$
$$\equiv \int \Delta^2(k) \mathrm{d} \ln k, \tag{1.48}$$

where the $\Delta^2(k) \equiv (2/\pi^2)k^3 P(k)$ represents the contribution to the variance per $\ln k$ interval. This quantity is useful as it does not depend on the Fourier transform convention used in the definition of the power spectrum, and relates more directly to intuitive statistics such the variance of randomly placed cells.

See *e.g.* Efstathiou (1996) for details of common methods of estimating these and other statistics.

1.4.2 Gaussian Random Fields

A particularly simple and important class of random fields are Gaussian Random Fields. These fields have the very important feature that their statistical properties are fully specified by the power spectrum $P(k)$ or, equivalently, the two-point correlation function $\xi(r)$. The probability distribution for $\delta(\mathbf{x})$ evaluated at n positions is an n-dimensional multivariate Gaussian

$$p(\delta_1, \ldots, \delta_n) \, \mathrm{d}\delta_1 \ldots \mathrm{d}\delta_n = \frac{1}{(2\pi)^{n/2} |\mathbf{C}|^{1/2}} \exp\left(\frac{1}{2} \sum_{ij} \delta_i \, C_{ij}^{-1} \delta_j \right) \mathrm{d}\delta_1 \ldots \mathrm{d}\delta_n, \tag{1.49}$$

where $\delta_i \equiv \delta(\mathbf{x}_i)$ and the covariance matrix is $C_{ij} = \langle \delta_i \delta_j \rangle = \xi(r_{ij})$. In these fields the phases of the Fourier modes are random, and the amplitude is distributed with an exponential distribution

$$p(|\delta_\mathbf{k}|^2 > P) = \exp\left(-\frac{P}{\langle |\delta_\mathbf{k}|^2 \rangle} \right). \tag{1.50}$$

This distribution results from the real and imaginary parts of $\delta_\mathbf{k} = a_\mathbf{k} + i b_\mathbf{k}$ having independent Gaussian distributions as follows. The bivariate Gaussian probability distribution is

$$g_\mathbf{k}(a_\mathbf{k}, b_\mathbf{k}, t) \mathrm{d}a_\mathbf{k} \mathrm{d}b_\mathbf{k} = \frac{1}{2\pi\sigma_\mathbf{k}^2(t)} \exp\left(-\frac{a_\mathbf{k}^2 + b_\mathbf{k}^2}{2\sigma_\mathbf{k}^2(t)} \right) \mathrm{d}a_\mathbf{k} \mathrm{d}b_\mathbf{k}, \tag{1.51}$$

where $\sigma_\mathbf{k}^2(t)$ is the variance in both $a_\mathbf{k}$ and $b_\mathbf{k}$. Again by homogeneity and isotropy, $\sigma_\mathbf{k}^2(t) = \sigma_k^2(t)$ depends only on the magnitude of $\mathbf{k}$.

If we write $\delta_{\mathbf{k}} = r_{\mathbf{k}} \exp(i\phi_{\mathbf{k}})$, the same probability distribution can also be expressed in terms of the amplitude $r_{\mathbf{k}} = |\delta_{\mathbf{k}}|$ and phase angle $\phi_{\mathbf{k}}$ as

$$g_{\mathbf{k}}(r_{\mathbf{k}}, \phi_{\mathbf{k}}, t) \mathrm{d}r_{\mathbf{k}} \mathrm{d}\phi_{\mathbf{k}} = \frac{1}{2\pi\sigma_k^2} \exp\left(-\frac{1}{2}\frac{r_{\mathbf{k}}^2}{\sigma_k^2}\right) r_{\mathbf{k}} \mathrm{d}r_{\mathbf{k}} \mathrm{d}\phi_{\mathbf{k}}. \tag{1.52}$$

Integrating over the argand plane:

$$p(|\delta_{\mathbf{k}}|^2 > P) = \int_{\sqrt{P}}^{\infty} \frac{1}{\sigma_k^2} \exp\left(-\frac{1}{2}\frac{r_{\mathbf{k}}^2}{\sigma_k^2}\right) r_{\mathbf{k}} \mathrm{d}r_{\mathbf{k}}, \tag{1.53}$$

leading to equation (1.50).

Gaussian Random Fields are a prediction of most inflationary models, see Coles & Lucchin (1995) and Peebles (1993). Because of the Central Limit Theorem (*e.g.* Kenney & Keeping 1951), any field with random phases will obey equation (1.49) – the sum of many independent variables (the contributions from each k-mode) has a Gaussian distribution. This should also work for the observed power spectrum, which is a convolution of the true power – see *e.g.* Peacock (1992). Because of this, the random phase condition alone is often used as a definition of a Gaussian Field. The distribution of powers has been directly measured by Feldman, Kaiser & Peacock (1994) and was found to obey equation (1.50), although this does not prove Gaussianity – see Fan & Bardeen (1995).

1.4.3 The Primordial Origin of Structure

There are two main candidates for the initial fluctuations in the universe, topological defects and quantum fluctuations blown up by inflation.

The more popular scenario is where quantum fluctuations in the early universe are blown up by inflation to scales greater than the Hubble length, and so are 'frozen in' until the Hubble length has expanded beyond their length scale. Reviews of this scenario can be found in Brandenberger (1990) and Kolb & Turner (1990), but there are several general predictions. The primordial power spectrum is expected to be close to the *Harrison-Zeldovich* spectrum, $P(k) \propto k$. This spectrum is called *scale invariant* because the *r.m.s.* potential fluctuations are the same on all scales at a given time, and the *r.m.s.* density fluctuations are all equal at the time they enter the horizon (*i.e.*

their lengthscales come within the Hubble length). This prediction arises essentially because of the time symmetry of de Sitter space, so that perturbations at the Hubble length (just about to leave the Horizon) must be the same for all t, by symmetry. The evolution of the inflaton can slightly alter this prediction and give a 'tilt' to the slope of the spectrum – see *e.g.* Liddle & Lyth (1992). An important prediction of most inflationary models is that the density field will be a *Gaussian Random Field*, see section (1.4.2). Perturbations produced during inflation are also usually expected to be adiabatic, *i.e.* have matching fluctuations in matter and radiation, see Efstathiou (1990).

Topological defects are believed to have formed during phase transitions in the early universe. There are several different types, including cosmic strings and textures – see *e.g.* Vilenkin & Shellard (1994). Their effect on the density field is not well understood, but in general they are expected to produce a field which is non-Gaussian.

1.4.4 The Pre-Recombination Era – The Transfer Function

Before recombination at $z \sim 1000$ the evolution of the density field depends on a combination of baryonic matter, dark matter and photons which evolve under self gravity, pressure and dissipative effects. This epoch must be modelled numerically (see *e.g.* Efstathiou 1990), but a few general insights are possible; in general the primordial spectrum of perturbations is altered, with small scales suppressed relative to large scale modes. These effects are characterised by a transfer function $T(k)$, the ratio of the late-time power spectrum to the primordial power $P_i(k)$,

$$P(k) = T^2(k)P_i(k). \tag{1.54}$$

The details of $T(k)$ depend on the constituents of the universe. If the dark matter is hot, *i.e.* fast (relativistic), density perturbations are very heavily suppressed because the particle motions smear out the structure, a process known as 'free streaming'. The cut off in the spectrum is the scale of the Horizon at the epoch when the particles become non-relativistic and the free streaming lengthscale drops. Perturbations in a cold dark matter model retain far more small scale power, but there is still a change of

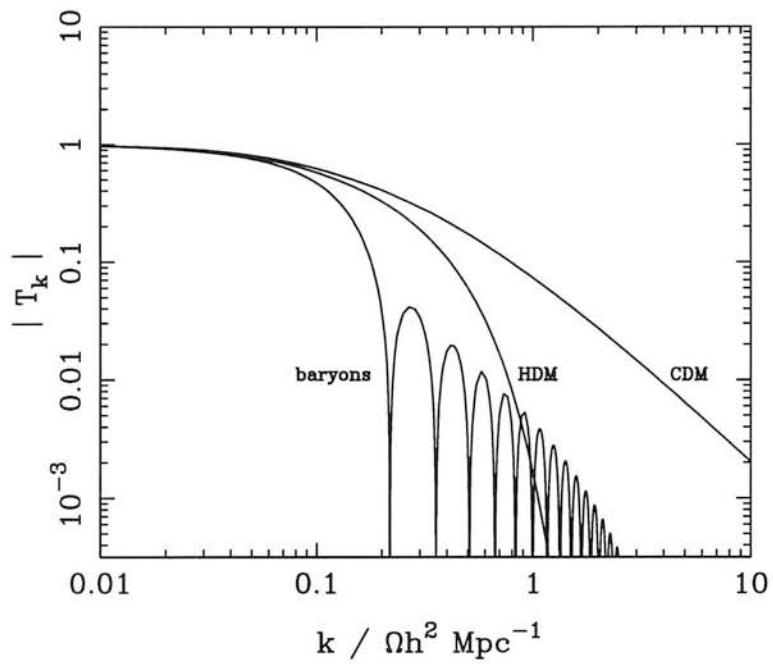

Figure 1.1: Transfer functions for example adiabatic models. The baryonic model does not scale exactly with Ωh^2, the case here is for $\Omega = 1$, $h = 0.5$.

slope. The evolution of perturbations in a radiation dominated phase is different outside the horizon compared with inside it – within the horizon the fluctuation growth is frozen. Perturbations entering the horizon before matter-radiation equality are frozen while larger ones still grow, changing the shape of the spectrum. In the matter dominated phase both growth rates are the same so the spectral shape is fixed again until the growth becomes nonlinear. The turnover in the spectrum is at the scale which enters the horizon at $t_{\rm equality}$. If the dominant matter is in fact baryonic, the situation is more complicated. Photons couple to the plasma, providing pressure support, so that overdensities within the horizon oscillate rather than grow up until recombination (rather than equality). The wiggles in the transfer function reflect the phase of these oscillations at recombination as a function of scale. There is an additional effect called Silk damping (Silk 1967), where photons begin to diffuse out of potential wells as decoupling approaches, and drag some of the matter with them. See Efstathiou (1990) for details of these effects.

Example fitting formulae from Bardeen *et al.* (1986) are, for CDM:

$$T(k) = \frac{\ln(1+2.34q)}{2.34q}[1+3.89q+(16.1q)^2+(5.46q)^3+(6.71q)^4]^{-1/4}; \qquad (1.55)$$

and for HDM:

$$T(k) = \exp(-3.9q - 2.1q^2), \qquad (1.56)$$

where $q = k/(\Omega h^2)\mathrm{Mpc}^{-1}$. These curves are illustrated in fig 1.1, along with a baryonic example for which there is no simple formula (see Efstathiou 1990). Curves for CDM models are often parameterised by $\Gamma \equiv \Omega h$.

1.4.5 Growth of Perturbations under Gravity

Gravitational instability tends to make perturbations grow with time. This process is usually divided into two parts; the linear regime where fluctuations are small and the behaviour is simple and well behaved, and the more complex nonlinear regime of large overdensities ($\delta > 1$) where various approximations are used. For scales smaller than the Hubble length, Newtonian gravitational theory is adequate to describe the evolution of clustering (see Peebles 1980, §6).

The following analysis applies to large scales where non-gravitational effects such as pressure are negligible.

Fluid Equations in Cosmology

The matter in the universe is modelled as a self-gravitating, pressureless fluid

$$\begin{aligned} \left(\frac{\partial \rho}{\partial t}\right)_{\mathbf{r}} + \nabla_{\mathbf{r}} \cdot \rho \mathbf{u} &= 0, \\ \left(\frac{\partial \mathbf{u}}{\partial t}\right)_{\mathbf{r}} + (\mathbf{u} \cdot \nabla_r)\,\mathbf{u} &= -\nabla_{\mathbf{r}} \Phi_{\mathrm{grav}} \\ \nabla^2_{\mathbf{r}} \Phi_{\mathrm{grav}} &= 4\pi G \rho. \end{aligned} \qquad (1.57)$$

The subscript $\mathbf{r}$ indicates that the coordinates are in terms of the proper distance $\mathbf{r}$ (relative to a fixed origin), while $\mathbf{u}$ is the proper velocity. Φ_{grav} is the gravitational

potential. These are the continuity and Euler equations which describe conservation of mass and momentum respectively, and also the Poisson equation for the Newtonian gravitational potential.

In an expanding universe it is convenient to factor out the expansion and work with comoving coordinates which expand with the scalefactor $a(t)$. The displacement $\mathbf{r}$ is related to the comoving equivalent $\mathbf{x}$ by $\mathbf{r} = a(t)\mathbf{x}$ and $\mathbf{u}$ is related to the peculiar velocity $\mathbf{v}$ by

$$\mathbf{u} = \dot{\mathbf{r}} = \dot{a}\mathbf{x} + \mathbf{v}(\mathbf{x}, t). \tag{1.58}$$

For example, the first term of the continuity equation becomes

$$\left(\frac{\partial}{\partial t}\right)_{\mathbf{r}} \rho(\mathbf{r}/a(t), t) = \frac{\partial \rho}{\partial t} - \frac{\dot{a}}{a}\mathbf{x} \cdot \nabla\rho, \tag{1.59}$$

while the divergence becomes

$$\nabla_{\mathbf{r}} \cdot \rho\mathbf{u} = \frac{1}{a}\nabla \cdot \rho\mathbf{v} + \frac{3\dot{a}}{a}\rho + \frac{\dot{a}}{a}\mathbf{x} \cdot \nabla\rho; \tag{1.60}$$

hence the new comoving continuity equation is given by

$$\frac{\partial \rho}{\partial t} + \frac{3\dot{a}}{a}\rho + \frac{1}{a}\nabla \cdot \rho\mathbf{v} = 0. \tag{1.61}$$

A similar transformation can be performed on the Euler and Poisson equations, leading to a system of comoving fluid equations,

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla \cdot (1+\delta)\,\mathbf{v} = 0, \tag{1.62}$$

$$\frac{\partial \mathbf{v}}{\partial t} + \frac{1}{a}(\mathbf{v} \cdot \nabla)\,\mathbf{v} + \frac{\dot{a}}{a}\mathbf{v} = -\frac{1}{a}\nabla\phi_{\text{grav}}, \tag{1.63}$$

$$\nabla^2\phi_{\text{grav}} = 4\pi G a^2 \bar{\rho}\delta \tag{1.64}$$

where δ is the dimensionless density contrast defined in equation (1.38) and the potential Φ_{grav} has been replaced with ϕ_{grav}:

$$\Phi_{\text{grav}} \equiv \phi_{\text{grav}} - \frac{1}{2}a\ddot{a}x^2. \tag{1.65}$$

Using the definition of the density parameter $\Omega = 8\pi G\bar{\rho}/(2H^2)$, the Poisson equation can be expressed in terms of Ω,

$$\nabla^2\phi_{\text{grav}} = \frac{3}{2}H^2 a^2 \Omega\delta \tag{1.66}$$

To calculate the evolution of the power spectrum, it is best to Fourier transform equations (1.62) to (1.64), to deal with the density field mode by mode,

$$\begin{aligned}
\frac{\mathrm{d}\delta_{\mathbf{k}}}{\mathrm{d}t} + \frac{1}{a} i\mathbf{k}\cdot\mathbf{v}_{\mathbf{k}} + \frac{1}{a}\sum_{\mathbf{k}'} i\delta_{\mathbf{k}'}(\mathbf{k}\cdot\mathbf{v}_{\mathbf{k}-\mathbf{k}'}) &= 0, \\
\frac{\mathrm{d}\mathbf{v}_{\mathbf{k}}}{\mathrm{d}t} + \frac{1}{a}\sum_{\mathbf{k}'} i\left[\mathbf{v}_{\mathbf{k}'}\cdot(\mathbf{k}-\mathbf{k}')\right]\mathbf{v}_{\mathbf{k}-\mathbf{k}'} + \frac{\dot{a}}{a}\mathbf{v}_{\mathbf{k}} &= i\frac{\mathbf{k}}{a}\phi_{\mathbf{k}}, \\
\mathbf{k}^2\phi_{\mathbf{k}} &= -4\pi G a^2 \bar{\rho}\delta_{\mathbf{k}}. \qquad (1.67)
\end{aligned}$$

Note that only the terms nonlinear in $\delta_{\mathbf{k}}$ and $\mathbf{v}_{\mathbf{k}}$ couple density fluctuations of different scales. Hence if the equations can be linearised, the modes grow independently.

Linear Growth of Perturbation

Analytic solutions for the growth of structure are possible when the density contrast is small, and the fluid equations can be expanded to first order in perturbation theory. In this simple *linear regime*, all overdensities grow at a constant rate, so the shape of the power spectrum $P(k)$ is fixed, only its amplitude changes.

In the linear perturbation approximation ($\delta \ll 1$), the perturbation equations become (by multiplying equation (1.62) by $\mathbf{v}$ and (1.63) by ρ and taking the divergence of the sum of both, see Peebles 1980)

$$\frac{\partial^2\delta}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial\delta}{\partial t} = 4\pi G\bar{\rho}\delta. \qquad (1.68)$$

For a flat ($\Omega = 1$), matter-dominated universe where $a \propto t^{2/3}$ and $6\pi G\bar{\rho}t^2 = 1$, equation (1.68) becomes

$$\frac{\partial^2\delta}{\partial t^2} + \frac{4}{3t}\frac{\partial\delta}{\partial t} = \frac{2}{3t^2}\delta \qquad (1.69)$$

which has a growing and a decaying solution

$$\delta = A(\mathbf{x})t^{2/3} + B(\mathbf{x})t^{-1}, \qquad (1.70)$$

Following the growing solution, the overdensity evolves as a power law as opposed to an exponential growth from the normal Jeans analysis in a static regime.

In general, the solution depends on the value of Ω (Peebles 1980). For $\Omega = 0$,

$$\delta = A(\mathbf{x}) + B(\mathbf{x})t^{-1}. \qquad (1.71)$$

Hence linear structure growth is frozen in an empty universe, although this result can be misleading; there is moderate growth ($\sim 1/3$ of the Einstein-de Sitter case) for $\Omega = 0.1$. Also, unless there is a significant vacuum contribution, Ω evolves away from unity – see equation (1.29). Hence an open universe may have had more rapid growth of linearly evolving structure in the past.

Similar calculations can be performed in Fourier space, leading to an equation essentially identical to (1.68),

$$\frac{\mathrm{d}^2\delta_{\mathbf{k}}}{\mathrm{d}t^2} + 2\frac{\dot{a}}{a}\frac{\mathrm{d}\delta_{\mathbf{k}}}{\mathrm{d}t} = 4\pi G\bar{\rho}\delta_{\mathbf{k}}, \tag{1.72}$$

with corresponding growing and decaying solutions. Clearly this equation also governs the growth of the power spectrum $P(k) \propto \langle|\delta_{\mathbf{k}}|^2\rangle$, which grows as the square of $\delta_{\mathbf{k}}$. The important point here is that $P(k)$ grows independently of k – the power spectrum retains its shape during linear growth.

There is much evidence that linear theory can be applied to the growth of large scales modes even if the density field has become highly nonlinear on small scales, although this has never been established rigorously. A plausibility argument is given by Peebles (1980, §28), the result following essentially because of momentum conservation – the centres of mass of large regions stay in position even if the particles move around on small scales. N-body experiments from Peebles & Groth (1976) to Peacock & Dodds (1994) seem to confirm this argument. The reliability of large-scale linearity is vital so that the shape of the pre-recombination power spectrum is retained, and the large scale velocity fields also obey simple linear equations.

Velocity Fields

The expansion of the universe greatly simplifies the analysis of the velocity field by removing vorticity. Consider the (linearised) comoving Euler equation – cf. equation (1.63),

$$\frac{\partial\mathbf{v}}{\partial t} + \frac{\dot{a}}{a}\mathbf{v} = -\frac{1}{a}\nabla\phi_{\mathrm{grav}}. \tag{1.73}$$

Taking the curl of both sides, the gradient vanishes and the remaining equation for $\nabla\times\mathbf{v}$ has a decaying solution $\propto 1/a$. Hence any vortical component of $\mathbf{v}$ will decay with time,

leading to an irrotational velocity field which can be expressed as the divergence of a scalar potential,

$$\mathbf{v} = -\nabla\Phi_{\mathbf{v}}. \tag{1.74}$$

The velocity field is still expected to be irrotational in the quasi-linear regime – see Dekel, Bertschinger & Faber (1990) for a detailed discussion. This assumption is vital for the POTENT method, discussed in section (1.5.1). The linearised version of the equation (1.62), the continuity equation, reads

$$\frac{\partial\delta}{\partial t} = -\frac{1}{a}\nabla\cdot\mathbf{v}. \tag{1.75}$$

Expressing the general linear growing-mode solution as $\delta = A(\mathbf{x})D(t)$, the continuity equation becomes (also using equation (1.74))

$$-\nabla^2\Phi_{\mathbf{v}} = \nabla\cdot\mathbf{v} = -a\delta\frac{\dot{D}}{D} = -a\delta\frac{f(\Omega)}{H}, \tag{1.76}$$

where $H = \dot{a}/a$ is the Hubble parameter, and the linear growth factor f is defined as

$$f = \frac{1}{H}\frac{\dot{D}}{D}. \tag{1.77}$$

This factor can be approximated quite accurately as $f \simeq \Omega^{0.6}$, see Peebles (1980, §14). Hence measurement of f from the velocity field or from the evolution of clustering gives a value for Ω. However, this assumes that the observed clustering of galaxies or clusters faithfully traces the underlying density field.

The relation between the potentials can be found from combining equations (1.76) and (1.66):

$$\phi_{\mathrm{grav}} = \frac{3Ha\Omega}{2f}\Phi_{\mathbf{v}}, \tag{1.78}$$

leading to the expression

$$\mathbf{v} = -\frac{3}{2}H^2a\Omega^{-0.4}\nabla\phi_{\mathrm{grav}} \tag{1.79}$$

for the velocity in terms of the potential.

1.4.6 Bias

Most observations of large-scale structure involve the analysis of the clustering of tracers, in particular galaxies. However, these objects may not faithfully trace the underlying distribution of (possibly non-baryonic) matter. The simplest possible assumption

is that of *linear bias*, where the galaxy and matter power spectra are related by

$$P^{\rm g}(k) = b^2 P^{\rho}(k) \tag{1.80}$$

where b is called the bias parameter. A similar but slightly more restrictive relation is

$$\delta^{\rm g}_{\mathbf{k}} = b\delta^{\rho}_{\mathbf{k}} \tag{1.81}$$

where the phases of the matter and galaxy modes are the same (unless b is allowed to be complex). Similarly in real space

$$\frac{\delta n}{n} = b\frac{\delta\rho}{\rho}, \tag{1.82}$$

although care must be taken with this definition as bias is expected to become non-linear below a certain scale.

Bias clearly exists – Peacock & Dodds (1994) compared the power spectra of several different tracers, *e.g.* optically selected galaxies and IRAS selected galaxies, and found them to have different amplitudes. Clearly, if one of the tracers is unbiased, the others won't be. The relative bias between tracers looks linear on large scales, providing some optimism about the assumption of linear bias relative to matter. The bias may be nonlinear, *i.e.* a function of scale, on smaller scales see *e.g.* Mann, Peacock & Heavens (1997) and Peacock (1997).

Under the assumption of linear bias as expressed in equation (1.81), the linearised continuity equation (1.75) becomes

$$\nabla \cdot \mathbf{v} = -a\delta\frac{\dot{D}}{D} = -a\delta_{\rm g}\frac{\beta}{H}, \tag{1.83}$$

where $\beta \equiv f/b$. Linear theory can only be used to constrain β, not Ω, because $\delta\rho/\rho$ cannot be measured directly.

See Dekel & Rees (1987) for a discussion of possible bias mechanisms.

1.5 Measuring Ω_0 with Galaxy Surveys

A number of techniques are used to try to constrain the density parameter Ω, usually via β, from surveys of galaxies. Much of this thesis uses the effect of redshift distortions to

constrain β – see chapter 2 and Hamilton (1997). Other popular methods for estimating β or Ω_0 from large-scale-structure data are discussed briefly here; see Dekel (1994) and Strauss & Willick (1995) for detailed reviews, and also Dekel, Burstein & White (1997) for a more general overview of measurements of Ω_0. Measures of Ω_0 from clusters are discussed briefly in section (1.3).

Note that with any of these methods, including the redshift distortion techniques, the parameter β is only equal to Ω/b if the velocity field is caused purely by gravitational instability. This may seem a reasonable assumption, non-gravitational initial velocities will decay as $\mathbf{v} \propto 1/a$ (see above), but the assumption has been questioned – see *e.g.* Babul *et al.* (1994).

1.5.1 Direct Peculiar Velocity Measurements and POTENT

If galaxy distances can be measured using a method other than redshift, the radial components of the peculiar velocities of the galaxies can be measured, by subtracting the 'Hubble flow velocity' from the velocity inferred from the redshift – see equation (2.2). These distance methods include the Tully-Fisher (TF) technique for spirals, and the D_n-σ method for ellipticals – see *e.g.* Jacoby *et al.* (1992). Bertcschinger & Dekel (1989) devised a method, known as POTENT, to take advantage the fact that the velocity field has a scalar potential (see equation (1.74)). The entire 3-dimensional velocity field can in principle be found by mapping potential at every point using only radial velocity information,

$$\Phi_{\mathbf{v}}(r_2, \theta, \phi) - \Phi_{\mathbf{v}}(r_1, \theta, \phi) = -\int_{r_1}^{r_2} v_r(r', \theta, \phi)\, \mathrm{d}r'. \tag{1.84}$$

The density and velocity fields can be compared using equation (1.83) to derive β. This method is discussed at great length in Dekel (1994). In a recent analysis using this method Sigad *et al.* (1997) found $\beta_{IRAS} = 0.89 \pm 0.12$.

The velocity field must be smoothed to try to make it as linear as possible, although the linear-nonlinear division appears to be sharpest in Fourier space. Smoothing reduces the number of independent points, and the surveys with velocity information are small and piecemeal. The errors on these distance measures are proportional to the distances

themselves, so these techniques can only be used for the local universe.

The TF and D_n-σ methods suffer from Malmquist bias, where only the brightest objects are seen at large distance. Some more recent variants of POTENT exist for reducing this, but they can introduce additional problems; see Dekel (1997) for a recent review.

1.5.2 The Dipole

The local group of galaxies is moving at about 600kms^{-1} relative to the microwave background radiation. By comparing this motion with the surrounding local density field which caused it (assuming gravitational instability), a limit on β can in principle be found using (cf. equation (1.83))

$$\mathbf{v} = \frac{H_0}{4\pi}\beta \int \delta \frac{\hat{\mathbf{r}}}{r^2}\, \mathrm{d}r^3. \tag{1.85}$$

This method is known as the Dipole technique and has been used by a number of authors.

Redshifts are actually unnecessary for this method; both gravity and light intensity obey an inverse-square law and thus the optical and gravitational dipoles should align. The method requires very complete sky coverage, and so is particularly well suited to use with sample drawn from the near-all-sky *IRAS* Point Source Catalogue (PSC). Dipole studies using the PSC include Meiksin & Davis (1986), and Yahil, Walker & Rowan-Robinson (1986); optical galaxies have been investigated by *e.g.* Lynden-Bell, Lahav & Burstein (1989).

The dipole requires a spherical survey of galaxy positions, and assumes that no overdensities beyond the range of the survey have any significant effect. Large-scale structure beyond the limit of the survey can add to the velocity. In practice the apparent value of β is measured as a function of an upper cutoff in redshift or magnitude until convergence is achieved. However, Peacock (1992) showed that a 'fake' convergence can easily occur, and that results in the literature could easily have errors which are considerably higher than those quoted by the authors. This is particularly misleading

because the direction of the gravitational dipole tends to align with the CMB (velocity) direction even when the apparent value of β has not converged.

Redshift surveys provide more information about the density field, and should therefore provide a more accurate estimate of β. *IRAS* redshift surveys are analysed in this way by Rowan-Robinson *et al.* (1990) and Strauss *et al.* (1992), finding $\beta_{IRAS} = 0.82 \pm 0.15$ and $\beta_{IRAS} = 0.55^{+0.20}_{-0.12}$ respectively. The latter paper includes a maximum likelihood analysis to account for unknown structure beyond the limit of the survey (although this is model-dependent). When redshifts are used, there is the problem of redshift distortions (see chapter 2). The observers motion actually introduces apparent structure in the observed density field, which the observer may assume is causing the motion in the first place – the 'Rocket Effect' – see Kaiser (1987).

Redshift distortion effects may themselves be used to derive information about the velocity field and hence Ω. This is explored in detail in chapter 2.

1.6 Thesis Overview

The Hot Big Bang model has been introduced, in particular with reference to the density parameter Ω and the origin and evolution of structure. In particular, statistics for measuring structure and some details of the linear growth of cosmological density and velocity fields were shown. Methods for measuring Ω were also briefly reviewed.

In chapter 2, the effects of redshift distortions caused by peculiar velocities on the measured clustering pattern are discussed both qualitatively and quantitatively. Power spectrum models for linear infall (Kaiser 1987) and nonlinear random velocities are introduced, and alternatives to the latter are considered. The linear-theory β estimator of Cole, Fisher & Weinberg (1994) is examined and extended to include the effects of nonlinear distortions. It is concluded that nonlinear distortion models are useful, but only in the nonlinearity is mild – clearly the largest scales must be analysed in present surveys to reliably measure β.

In chapter 3 the possibility of measuring a cosmological constant Λ from anisotropies in the measured clustering in redshift surveys of galaxies or quasars is examined. This effect is independent of the evolution of the sources. It is shown that the redshift distortions can mimic this effect, making it much more difficult to detect. A power spectrum model is introduced, which includes the effects of shot noise, linear and nonlinear redshift distortions as well as the geometric effect caused by Λ. A Likelihood estimator is used to test if the next generation of redshift surveys would be capable of constraining Λ, and it is found that a marginal constraint may be possible. This work is published as Ballinger, Peacock & Heavens (1996).

The need for analysing very large scale modes was discussed in chapter 2. In chapters 4 and 5, a spherical harmonic and spherical Bessel function transform is discussed and applied, following the work of Heavens & Taylor (1995). The real-space power spectrum of the 1.2Jy survey and the PSCz survey are derived along with simultaneous estimates of β for each survey. The 1.2Jy results are also published in Ballinger, Heavens & Taylor (1995).

In chapter 6 a method for linear data compression is introduced and tested. Initially, the Karhunen-Loève/SVD method of Tegmark, Taylor & Heavens (1997) was tested and found to have difficulty with highly correlated parameters. An alternative, 'Parameter Eigenvector Optimisation', is then introduced and found to be much more successful.

Chapter 7 includes a review of the results of this thesis, and suggestions for refinements and extensions to this work.

Chapter 2

REDSHIFT-SPACE DISTORTIONS

In recent years, with the advent of deep surveys of galaxies with accurate measurements of redshifts it has become possible to create three dimensional maps of the galaxy distribution at very large scales in the universe. These surveys can be used to derive information about the properties and structure of the universe, which can in turn give clues as to the properties of the very early universe. However, galaxy positions are measured using their redshifts as distance indicators, *i.e.* they are measured in *redshift space*, where peculiar velocities distort the radial coordinates.

Other measurements of distance exist, but they have been used for only a small subset of the large number of galaxy redshifts, they are prone to systematic uncertainty, and the distance errors increase linearly with distance, limiting their use to $r < 50\,h^{-1}\,\mathrm{Mpc}$ – see the discussion of the POTENT method in section (1.5).

If certain assumptions about the density and velocity fields are made, redshift distortions can actually be used to derive information about the velocities and hence the density parameter Ω. In particular, the density field is believed to be homogeneous and isotropic, so anisotropies in the observed redshift space field give information about the velocity field. This effect is statistical as the velocity of each individual galaxy is

unknown.

Recapping on the origin of peculiar velocities, consider the relation between proper and comoving coordinates

$$\mathbf{r} = a(t)\mathbf{x} \tag{2.1}$$

Here $\mathbf{r}$ is the position vector of a galaxy at time t, $a(t)$ is the expansion factor of the universe and $\mathbf{x}$ is the comoving coordinate, the position vector at time $t = t_0$ when $a(t_0) = 1$. Equation (2.1) can be differentiated with respect to time to give:

$$\dot{\mathbf{r}} = \dot{a}(t)\mathbf{x} + a(t)\dot{\mathbf{x}} \tag{2.2}$$

Equation (2) describes the motion of a galaxy in the universe. There are two components on the right hand side, the first is the *Hubble flow* which contains the true distance information and the second is the *peculiar velocity* $\mathbf{v} = a(t)\dot{\mathbf{x}}$ which distorts the distance information. If the peculiar velocity of a particle was zero its comoving coordinate $\mathbf{x}$ would be constant. The redshift distance is assigned by $\mathbf{s} = \dot{\mathbf{r}}/H_0$, which includes the peculiar velocity component.

This redshift space distortion effect leads to many problems as it alters the statistical properties of the density field. However, it also contains information about the statistical properties of the peculiar velocity field, which can be exploited to derive an estimate of $\beta \equiv f(\Omega)/b$, where $f(\Omega) \approx \Omega^{0.6}$, the dimensionless velocity factor and b is the bias parameter – see chapter 1.

For a very detailed review covering a wider scope than this chapter see Hamilton (1997), and see also Strauss & Willick (1995).

The redshift distortion effect is purely radial; the position vector of a galaxy is distorted by the radial component of its peculiar velocity $u \equiv \mathbf{v}(\mathbf{r}) \cdot \hat{\mathbf{r}}/H_0$. Any changes to the statistical properties of the galaxy density field are also purely radial – hence the observed structure becomes statistically *anisotropic*. In a statistical analysis this anisotropy can be used to quantify the distortion effect and constrain the density parameter Ω_0, in the context of a peculiar velocity model. If the peculiar velocities are coherent, *e.g.* linear velocity fields caused by inhomogeneities, the distortion causes a

mapping from real-space coordinates (r, θ, ϕ) to redshift space (s, θ, ϕ):

$$(r, \theta, \phi) \rightarrow (s, \theta, \phi) = (r, \theta, \phi) + (u, 0, 0), \tag{2.3}$$

If the velocities are measured in the Local Group rest frame rather than the CMB frame, there is an extra correction for the bulk motion of the obsever,

$$s^{LG} = r + v - \hat{\mathbf{r}} \cdot \mathbf{v}, \tag{2.4}$$

see Hamilton (1997).

2.1 Qualitative Effects of Redshift Distortions on the density field

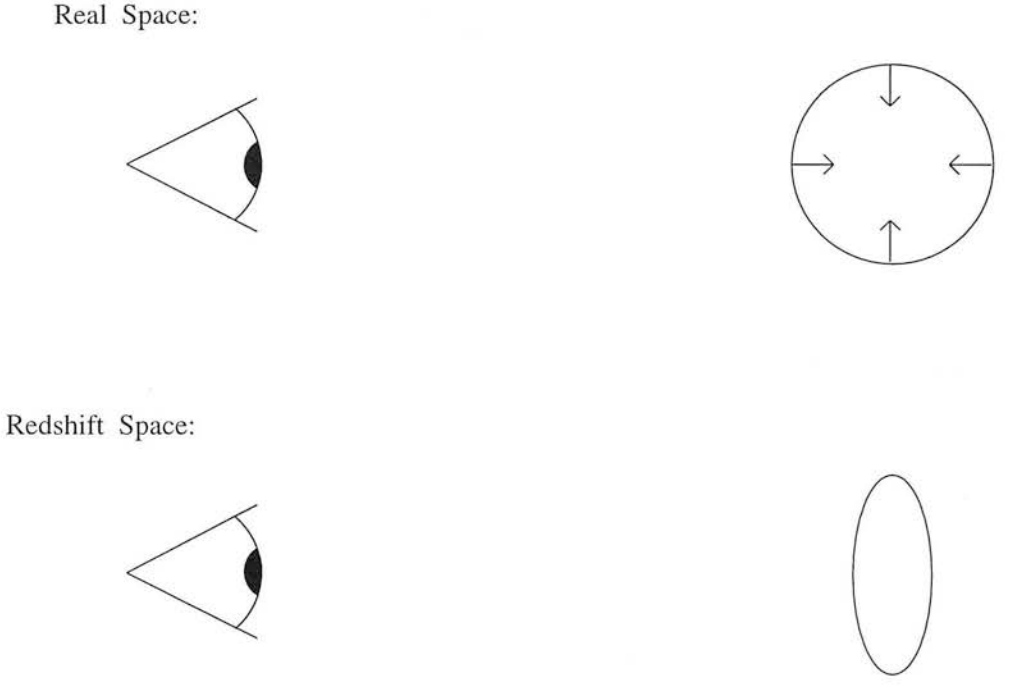

Figure 2.1: The compression of a large scale overdensity along the line of sight in redshift space. The arrows show the coherent infall caused by the overdensity.

The redshift distortions can have different effects on different scales. On very large scales, where peculiar velocities obey linear theory (because the fractional overdensities

are small, see Peebles (1980 §7)), coherent velocity fields due to the gravitational infall tend to compress overdense features along the line of sight – see Fig 2.1. This effect was originally considered by Sargent & Turner (1977). The distortion sharpens up the density contrast – overdense features look denser because more galaxies appear to be closer to their centres, underdense features seem even less dense because galaxies appear to be further from their centres. This effect can also be considered as an increase in the amplitude of radial waves – see Hamilton (1997).

Conversely, on smaller scales, incoherent motions of galaxies in virialised clusters cause the 'Finger of God' effect where clusters are stretched out along the line of sight in redshift space, see Fig 2.2. This effect was first pointed out by Jackson (1972).

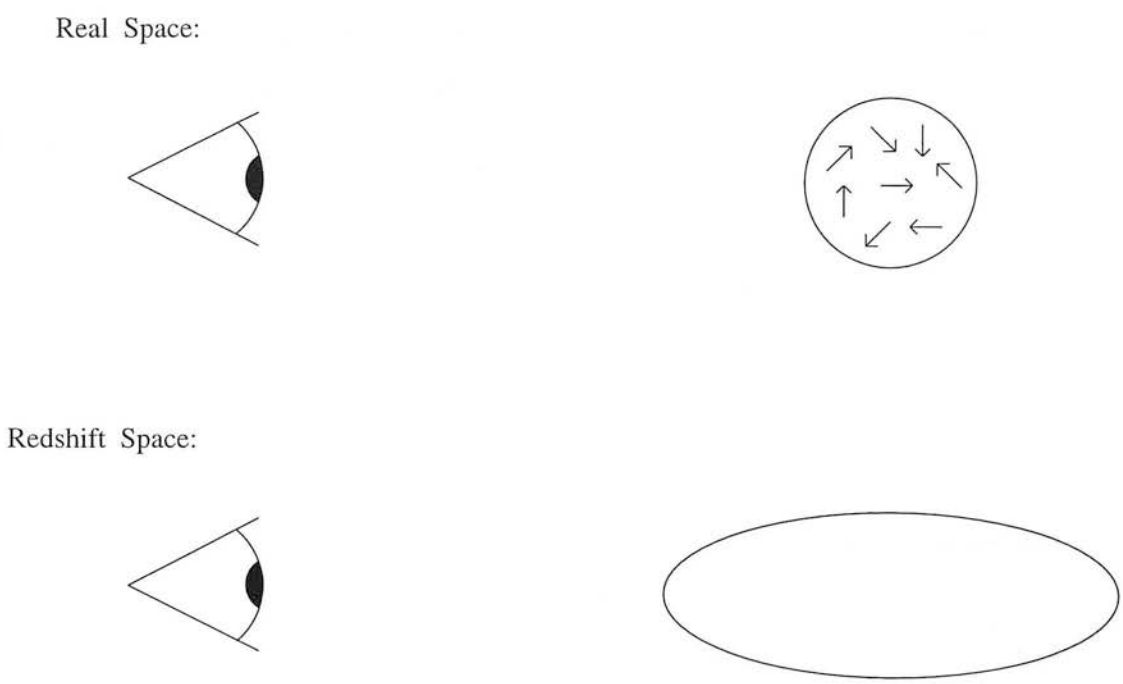

Figure 2.2: A 'Finger of God': the stretching of a virialised cluster along the line of sight in redshift space. The arrows denote the random motions of galaxies in the cluster.

The two distortion effects act in the opposite sense, coherent infall sharpens up radial structure, random motions blur it out. In general two regimes are expected, with enhanced structure at large scales and suppressed structure at small scales.

2.2 Redshift Distortions in the Linear Regime

Kaiser (1987) found a simple form for the effect of the large scale coherent velocity fields on the power spectrum assuming linear theory and the 'distant observer approximation'. This is the assumption that the observer is distant relative to the size of the survey volume so that all lines of sight can be treated as parallel. The Kaiser formula, derived below, is

$$P^S(k,\mu) = P^R(k)(1+\beta\mu^2)^2 \tag{2.5}$$

where superscripts R and S denote real and redshift space respectively, and μ denotes the cosine of the angle between the wavevector and the line of sight. It can be seen from the formula that the value of the power spectrum is boosted along the line of sight, fluctuations appearing stronger due to the increased density contrast. Also, it is anisotropic (it has angular dependence), with no effect perpendicular to the line of sight (when $\mu = 0$), as expected. This can be exploited to find an estimate of β from the observed redshift space power spectrum.

Kaiser's formula can be derived by analysing the effect of redshift distortions on a single density plane wave in the linear regime assuming the 'distant observer' approximation. This simplifies the analysis because the distortion direction can be taken one of the cartesian axes of the system.

Setting $H_0 = 1$, the mapping between real space ($\mathbf{r}$) and redshift space ($\mathbf{s}$) becomes

$$\mathbf{s}(\mathbf{r}) = \mathbf{r}\left[1 + \frac{u(\mathbf{r}) - u(0)}{r}\right], \tag{2.6}$$

leading to the Jacobian for the transformation of volume elements

$$\mathrm{d}^3 s = \mathrm{d}^3 r \left[1 + \frac{u(\mathbf{r}) - u(0)}{r}\right]^2 \left[1 + \frac{\partial u(\mathbf{r})}{\partial r}\right]. \tag{2.7}$$

It is best to perform this sort of redshift distortion analysis in the Local Group frame, so that the term $[u(\mathbf{r}) - u(0)]/r$ doesn't become large at small r. The first term corresponds to the change in the apparent physical size of an area of fixed angular size due to a change in apparent distance; the second term represents a change in the length of the volume along the line of sight. From linear theory, the peculiar velocity associated with a wave mode of wavenumber k and amplitude $\delta_k(\mathbf{k})$ is of order $u \sim \delta(\mathbf{k})/k$, while the

gradient is $du/dr \sim ku \sim \delta(\mathbf{k})$ – see section (1.4.5). For a distant observer, the survey volume V subtends a small angle so $V^{1/3}/r \ll 1$. Hence, for all modes smaller than the survey, $r \gg 1/k$, and the first term can be set to unity. The continuity condition

$$\rho(\mathbf{s})\mathrm{d}^3 s = \rho(\mathbf{r})\mathrm{d}^3 r \tag{2.8}$$

leads to

$$\rho^S(\mathbf{s}) = \rho^R(\mathbf{r}) \left[1 + \frac{\partial u(\mathbf{r})}{\partial r}\right]^{-1}, \tag{2.9}$$

which, to first order in $\partial u/\partial r$, yields

$$\rho^S(\mathbf{s}) = \rho^R(\mathbf{r}) \left[1 - \frac{\partial u(\mathbf{r})}{\partial r}\right]. \tag{2.10}$$

In terms of density contrast $\delta^R(\mathbf{r}) \equiv \rho(r)/\overline{\rho} - 1$ with $\delta^R(\mathbf{r}) \ll 1$ we find

$$\delta^S(\mathbf{s}) = \delta^R(\mathbf{r}) - \frac{\partial u(\mathbf{r})}{\partial r}. \tag{2.11}$$

Kaiser's formula is derived by considering a density field made up of a single plane wave,

$$\delta^R(\mathbf{r}) = \delta(\mathbf{k}) \cos(\mathbf{k} \cdot \mathbf{r} + \phi), \tag{2.12}$$

with the corresponding velocity from linear theory (see equation (1.83)),

$$\mathbf{v}(\mathbf{r}) = -\beta\, \delta(\mathbf{k}) \frac{\mathbf{k}}{k^2} \sin(\mathbf{k} \cdot \mathbf{r} + \phi). \tag{2.13}$$

Note that this calculation involves the density field of galaxies, not of mass, see section (1.4.6). The redshift distortion is caused by the radial projection of the velocity,

$$u(\mathbf{r}) = \mathbf{v}(\mathbf{r}) \cdot \hat{\mathbf{r}} = -\frac{\mu\beta}{k}\, \delta(\mathbf{k}) \sin(\mathbf{k} \cdot \mathbf{r} + \phi); \tag{2.14}$$

and differentiating to find the projected gradient gives

$$\frac{du}{dr} = -\mu^2 \beta\, \delta(\mathbf{k}) \cos(\mathbf{k} \cdot \mathbf{r} + \phi) = -\mu^2 \beta\, \delta^{\mathrm{R}}(\mathbf{r}). \tag{2.15}$$

Substituting for the projected velocity gradient in equation (2.11) gives

$$\delta^S(\mathbf{s}) = \delta^R(\mathbf{r})(1 + \mu^2 \beta). \tag{2.16}$$

This shows that a real-space plane wave maps to a redshift-space plane wave with the same phase but an amplitude enhancement which depends on μ, the cosine of the angle between the wavevector and the line of sight. For $\mu = 0$ we find that $\delta^S(\mathbf{s}) = \delta^R(\mathbf{r})$

as expected – perpendicular structure is unaffected. Clearly from equations (2.16) and (2.12)

$$\delta^S(\mathbf{k}) = \delta^R(\mathbf{k})(1 + \mu^2\beta); \tag{2.17}$$

ensemble averaging leads to the Kaiser equation for the power spectrum – equation (2.5):

$$P^S(k, \mu) = P^R(k)(1 + \mu^2\beta)^2.$$

The Kaiser equation clearly illustrates the statistical anisotropy of the redshift space density field. The effect can be seen in Fig 2.3. Note that the Kaiser formula for $P(k)$ only makes any sense if μ is independent of position. This is true in the distant observer limit.

Taking the azimuthal average of equation (2.5) gives the net boost to the angle-average power in a survey,

$$P^S(k) = P^R(k)\left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right). \tag{2.18}$$

For an unbiased Einstein-de Sitter universe ($\beta = 1$) this boosts the angle averaged power spectrum by a factor of 28/15.

The Kaiser equation is derived under very restrictive assumptions – distant observer, modes are independent – but it is widely applied to a large number of areas, particularly in the form of equation (2.18). A very much more detailed analysis was performed by Zaroubi & Hoffman (1996), who showed that for a general survey with a selection function $\phi(\mathbf{r})$ and underlying real-space power spectrum $\langle \delta_r(\mathbf{k}')\delta_r^*(\mathbf{k}'')\rangle = (2\pi)^3 P(k')\delta_{\rm D}(\mathbf{k}' - \mathbf{k}'')$ the (no longer diagonal) Fourier mode covariance function becomes

$$\langle \delta_s(\mathbf{k}_1)\delta_s^*(\mathbf{k}_2)\rangle = \left(\frac{1}{2\pi^2}\right)\int \mathrm{d}^3k' P(k')\left[I_0(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}') + \beta I_1(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}') + \beta^2 I_2(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}')\right], \tag{2.19}$$

where the convolution Kernels are given by

$$I_0(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}') = \left\{\int \exp[-i(\mathbf{k}_1 - \mathbf{k}')\cdot\mathbf{r}_1]\phi(\mathbf{r}_1)\mathrm{d}^3r_1\right\}\left\{\int \exp[i(\mathbf{k}_2 - \mathbf{k}')\cdot\mathbf{r}_2]\phi(\mathbf{r}_2)\mathrm{d}^3r_2\right\}$$

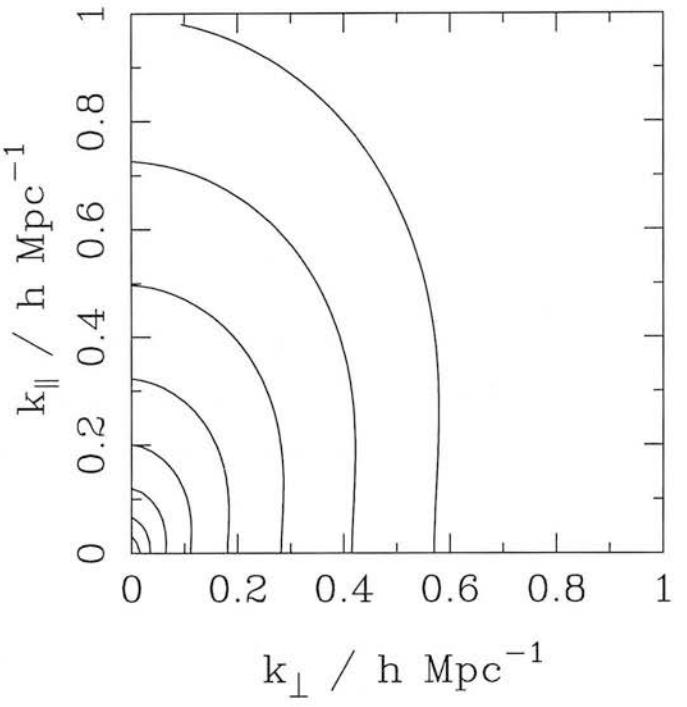

Figure 2.3: Power spectrum contours for the purely linear formula (equation (2.5)) with $\beta = 0.5$. The power spectrum is boosted along the line of sight, stretching the contours out in the $k_{||}$ direction. The real space contours would be quarter circles due to isotropy. The assumed real-space power spectrum is a simple power law $P(k) \propto k^{-1.5}$. Note that $P(k, \mu)$ has its highest value in the bottom left corner. The contours are spaced in powers of two.

$$
\begin{aligned}
I_1(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}') &= \hat{k}'_i \hat{k}'_j \left\{ \int \hat{\mathbf{r}}_{1,i} \hat{\mathbf{r}}_{1,j} \exp[-i(\mathbf{k}_1 - \mathbf{k}') \cdot \mathbf{r}_1] \phi(\mathbf{r}_1) \mathrm{d}^3 r_1 \right\} \\
&\times \left\{ \int \exp[i(\mathbf{k}_2 - \mathbf{k}') \cdot \mathbf{r}_2] \phi(\mathbf{r}_2) \mathrm{d}^3 r_2 \right\} \\
&+ \hat{k}'_i \hat{k}'_j \left\{ \int \exp[-i(\mathbf{k}_1 - \mathbf{k}') \cdot \mathbf{r}_1] \phi(\mathbf{r}_1) \mathrm{d}^3 r_1 \right\} \\
&\times \left\{ \int \hat{\mathbf{r}}_{2,i} \hat{\mathbf{r}}_{2,j} \exp[i(\mathbf{k}_2 - \mathbf{k}') \cdot \mathbf{r}_2] \phi(\mathbf{r}_2) \mathrm{d}^3 r_2 \right\} \\
I_2(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}') &= \hat{k}'_i \hat{k}'_j \hat{k}'_l \hat{k}'_m \left\{ \int \hat{\mathbf{r}}_{1,i} \hat{\mathbf{r}}_{1,j} \exp[-i(\mathbf{k}_1 - \mathbf{k}') \cdot \mathbf{r}_1] \phi(\mathbf{r}_1) \mathrm{d}^3 r_1 \right\} \\
&\times \left\{ \int \hat{\mathbf{r}}_{2,l} \hat{\mathbf{r}}_{2,m} \exp[i(\mathbf{k}_2 - \mathbf{k}') \cdot \mathbf{r}_2] \phi(\mathbf{r}_2) \mathrm{d}^3 r_2 \right\}
\end{aligned}
\tag{2.20}
$$

with summation implied. This complicated formula shows a number of features which the Kaiser formula lacks. Clearly there is a mode-mixing effect: the observed power is

a convolution over the underlying power. Kernels I_1 and I_2 represent additional mixing due to redshift distortions; this is on top of that caused by the selection function only (kernel I_0), and should be accounted for on large scales. This covariance matrix is in effect 6D as there are modes in three cartesian directions mixing with each other. A full likelihood analysis using this transform would involve the repeated inversion of this matrix, so the computational expense could be considerable. In chapter 4 an alternative, spherical transform is used which reduces the size of the matrix.

The linear redshift space correlation function can also be derived but the formalism is more messy, see Lilje & Efstathiou (1989) and Hamilton (1992). As mentioned in chapter 1, the division between the linear and nonlinear regimes seems clearer in Fourier space, so the correlation function would appear to be a less satisfactory statistic to use.

2.3 Non-Linear Redshift Distortions

Small scale distortions where the velocity field has become nonlinear are more difficult to treat accurately, but some approximations can be used. In general they involve a suppression of effective β, if β is measured assuming linear theory.

2.3.1 Incoherent Velocities – The Fingers of God

Peacock (1992) suggested that the 'Fingers of God' from random virialised motions could be approximated by an incoherent Gaussian scatter along the line of sight. This additional probability distribution is convolved with the correlation function which is equivalent to multiplying its Fourier transform (also a Gaussian) with the power spectrum:

$$P^S(k,\mu) = P^R(k) D_L \exp\left[-\frac{1}{2}\left(\frac{k\mu\sigma_{\rm p}}{H}\right)^2\right] \qquad (2.21)$$

where D_L is the linear factor as in equation (3). The term $\sigma_{\rm p}$ is the *pairwise* velocity dispersion of the galaxies (Peacock uses the one particle dispersion which is $\sqrt{2}$ times smaller). Again it depends on angle and has no effect perpendicular to the line of sight. This formula assumes that the velocity dispersion is the same for every galaxy,

which is clearly untrue. In the centres of clusters, where the 'Fingers of God' are large and obvious, the velocity dispersions can easily be over twice the average – see Mo, Jing & Börner (1993). Peebles (1976) and later Davis & Peebles (1983) suggested that observationally the pairwise velocity distribution is best fitted by an exponential form. This distribution has wider wings, perhaps because it is made up of a combination of Gaussians with different dispersions – see Sheth (1996). An exponential distribution is equivalent to multiplying the power spectrum by a Lorentzian:

$$P^S(k,\mu) = P^R(k) D_L \left[\frac{1}{1 + \frac{1}{2}\left(\frac{k\mu\sigma_{\rm p}}{H}\right)^2} \right] \tag{2.22}$$

It is actually rather unclear what $\sigma_{\rm p}$ really represents. The effect is modelled as a random scatter on the velocity of every galaxy, regardless of environment, but this clearly is not what happens in reality. Large cluster have velocity dispersions of $\sim$ 1000km s^{-1}, whereas the measured mean value is about 400km s^{-1}. Other non-linear effects can also cause the suppression of 'effective β', such as quasi-linear infall, where the overdensity:velocity ratio increases relative to the linear prediction.

In a simulation, where the velocity of every galaxy is known, the pairwise velocity dispersion is calculated simply by averaging the relative velocity of every pair of galaxies. This is clearly not exactly the same $\sigma_{\rm p}$ as appears in equations (2.21) and (2.22) as can be seen by considering a simulation where every galaxy is assigned its velocity according to linear theory. A non-zero velocity dispersion will be measured – with a Gaussian distribution if the density field is Gaussian – but the effective $\sigma_{\rm p}$ of 2.21 and 2.22 will be zero, Kaiser's formula will be (statistically) exact on all scales. With real data, $\sigma_{\rm p}$ is measured from correlation function anisotropies in a very uncertain way, see Mo *et al.* (1993). The scatter can also represent random redshift errors, which would add in quadrature to any underlying value of velocity scatter. Considering all of the uncertainties, it is perhaps best to fit $\sigma_{\rm p}$ as a free parameter, and not derive it from other methods *e.g.* the small scale correlation function.

Peebles (1976) and later Davis & Peebles (1983) did derive $\sigma_{\rm p}$ from the redshift space correlation function, and tried to measure Ω from the small scale effect directly. The physical model used was the *cosmic virial theorem*, and they assumed that all

the clustering creating the measured stretching in the correlation function is stable and virialised, although Davis & Peebles (1983) also used a simple infall model for intermediate scales. In the light of the previous considerations, and also the greater risk of nonlinear biasing on small scales, these results seem far less reliable than those built on linear theory. More recently, Jing, Mo & Börner (1997) suggested that the Las Campanas Redshift survey contained a fair sample of rich clusters, but their analysis depended very heavily on extensive modelling; see also Dekel, Burstein & White (1997). Fisher (1995) modified the streaming model of Davis & Peebles (1983) to be consistent with the Kaiser analysis at large scales.

An alternative nonlinear model was proposed by Taylor & Hamilton (1996) and Fisher & Nusser (1996). The proposal was that the infall analysis could be carried over into the quasi-linear regime using the Zeldovich approximation (Zeldovich 1970), which assumes linear growth of particle trajectories rather than overdensities. These models predict a fall in apparent β but for a very different reason. Essentially, as linearity breaks down, the density contrast rises faster than the velocity (relative to the linear prediction). If the effect is weak it may be modelled as an extra $\sigma_{\rm p}$ component, but this throws away extra information on β from the detailed infall modelling – but if this information is used, the model needs to be very reliable. It has been suggested that this model is a better approximation for galaxies which are under-represented in virialised clusters, such as *IRAS* selected galaxies.

A recent, detailed analysis (Hatton & Cole 1997) of both the random smearing and Zeldovich models suggests that neither of them are very accurate if the density field is substantially nonlinear. It would seem that ideally, nonlinear scales should be avoided altogether.

2.4 Estimating β from Power-Spectrum Anisotropies

A number of authors have constrained β using the azimuthal average of the power spectrum (or correlation function) in redshift space – equation (2.18) – see *e.g.* Peacock & Dodds (1994), Fry & Gaztañaga (1994) and Peacock (1997). This involves comparing

a redshift space power spectrum with a real space power spectrum, *e.g.* one derived by deprojecting angular clustering information (with no redshift information, see *e.g.* Baugh & Efstathiou 1993) in order to produce an estimate of β. However, Hamilton (1992,1993) developed a method for measuring β by observing the anisotropy of the correlation function, *i.e.* its angular dependence relative to the line of sight. Cole, Fisher & Weinberg (1994) (hereafter CFW94) recreated Hamilton's method in the neater power-spectrum formalism, where linearity can be made more reliable (see chapter 1) and also the error assignment is more straightforward and allows β errors to be obtained, see chapter 3, although this was not done properly by Cole, Fisher & Weinberg (1995) – see later. The technique involves estimating β from moments of the angular power spectrum assuming the Kaiser equation *i.e.* assuming linear theory. CFW94 examined the effects on this estimate due to various non-linear effects, including small scale velocity dispersions, using N-body simulations.

In this section the multipole formula for β is recalculated with the additional Finger-of-God formulae of equation (2.21) and equation (2.22). In a follow-up paper – Cole, Fisher & Weinberg (1995, CFW95) – a similar analysis was performed.

2.4.1 Linear theory power-spectrum multipoles

CFW94 developed a method for estimating β from a multipole expansion of the power spectrum, based on the correlation function analysis of Hamilton (1992). The anisotropic spectrum is decomposed into Legendre Polynomials (denoted $\mathcal{P}_l(\mu)$):

$$P^S(k,\mu) = \sum_{l=0}^{\infty} P_l^S(k)\mathcal{P}_l(\mu) \tag{2.23}$$

with multipole moments:

$$P_l^S(k) = \frac{2l+1}{2}\int_{-1}^{+1} \mathrm{d}\mu\, P^S(k,\mu)\mathcal{P}_l(\mu) \tag{2.24}$$

Using equation (2.5), the value of β can be recovered from the ratio of the quadrupole($l = 2$) to monopole($l = 0$) moments:

$$\frac{\mathcal{P}_2^S(k)}{\mathcal{P}_0^S(k)} = G(\beta) = \frac{\frac{4}{3}\beta + \frac{4}{7}\beta^2}{1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2} \tag{2.25}$$

This can be inverted to find β. This estimator seems extremely useful as it does not depend on the shape of the power spectrum as a function of k, merely on its angular dependence in redshift space. This method is based on the Kaiser equation and so depends on the distant observer approximation.

An important point must be made about parameter estimation and error analysis. The inversion using harmonic multipoles is very elegant but it merely gives a *number*, not a *constraint* – there is no error bar. A seemingly clumsier alternative is to use the known stochastic properties of the power spectrum and employ a Likelihood method to directly compare a model $P(k, \mu)$ with the data – see chapter 3. This approach has the advantage that the likelihood parameter space can be explored, giving full error bars on all the parameters, including β. In principle full errors from proper covariances can be calculated for the multipole method, although the formalism becomes rather elaborate; see de Laix & Starkman (1997). Instead CFW95 measure the covariances from a set of simulations all of which have $\beta = 0.5$. This is effectively calculating errors in reverse – taking the result from the multipole ratio and assuming that this is the true value. The error on β increases with β, so large values may be unfairly excluded.

2.4.2 Fingers of God

The above estimator depends upon Kaiser's linear theory redshift space effect. However, the β estimate may be affected by non-linear effects, in particular the 'Finger of God' effect (CFW94 refers to this as 'dispersion non-linearity'). This random dispersion effect acts in the opposite way to the linear effect, 'washing out' the density contrast by randomly moving galaxies around in redshift space. Hence it suppresses the apparent value of β derived from a linear theory estimator.

CFW94 remove the effect in their simulations by smoothing the particle velocities. This identifies the effect but cannot be done with real data where the actual velocities are unknown. The following analysis gives the effects on this β estimator of the two analytic FoG models given above. This allows the 'Finger of god' effect to be isolated from any other non-linear effect (CFW94 identify another three) and gives a clearer

understanding of the processes involved.

The results show that the two models have a very similar effect on fairly large scales, and are themselves similar in effect to the small scale distortions in the N-body studies of CFW94, as also demonstrated in CFW95.

It was found that the moments of the power spectrum (7) could still be calculated analytically with either of the two non-linear models included. This allowed the effect on the estimate of β to be calculated analytically. See appendix A for calculations and formulae for moments.

The pairwise velocity dispersion $\sigma_{\rm p}$ in the two models is empirical – calculated from observed small scale redshift space correlation functions, or ideally fitted as an extra free parameter. Its value is very uncertain, the measurement depending strongly on the richness of the survey from which it is derived, see Mo, Jing & Börner (1993). A plausible average value is:

$$\sigma_{\rm p} = 450 \mathrm{km\ s^{-1}}, \tag{2.26}$$

although the models were tested using several different values.

It was expected that at large enough scales the two models would look identical since any symmetric distribution of incoherent velocities would have the same form in k-space for small k. If $f(x)$ is a distribution in real space, then its Fourier transform is:

$$\begin{aligned} F(k) &= \int_{-\infty}^{\infty} f(x) \exp(ikx)\, \mathrm{d}x \\ &= \int_{-\infty}^{\infty} f(x) \left[1 + ikx - \frac{k^2x^2}{2} + O(k^3)\right] \mathrm{d}x \\ &= 1 - \frac{k^2\sigma^2}{2} + O(k^4) \end{aligned} \tag{2.27}$$

– for any symmetric, normalised distribution. Here we use the fact that $\int x^2 f(x)\, \mathrm{d}x = \sigma^2$, the variance of the real space distribution. Odd k terms vanish in the integral because they are antisymmetric.

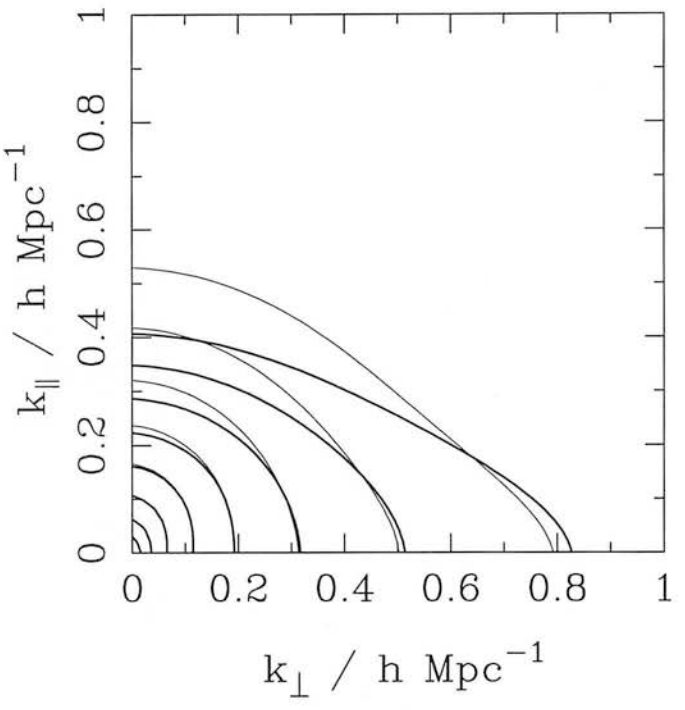

Figure 2.4: Contours of $P(k)$ with $\beta = 0.5$ including the effects of the 'Fingers of God' with $\sigma_p = 450\text{km s}^{-1}$. The thick contours are for the the Gaussian model (equation (2.21)), the thin contours for the exponential model (equation (2.22)). The power spectrum is suppressed along the line of sight at large k (small scale), squashing the contours out in the $k_{||}$ direction. At small k (bottom left corner) the linear boost effect dominates again.

2.4.3 Results

In order to compare our results with the N-body results of CFW94 they were plotted in the same way. To show the general anisotropic effect of the redshift distortion, contour plots of the power spectrum in the plane defined by $k_{||}$ and $k_{\perp}$, wavevector components parallel and perpendicular to the line of sight, are shown. Note that, in these contour plots, large scales are compressed into the bottom left corner. To show the effect of small scale dispersions on the CFW94 estimator the ratio of the estimated β to true β is plotted against wavelength (scale). In Fig 2.3, the purely linear boost effect is shown. Contours on all scales are boosted along the line of sight by the Kaiser effect

equation (2.5). The CFW94 technique assumes this effect. Note that $P(k, \mu)$ has its highest value in the bottom left corner. The contours are spaced in powers of two.

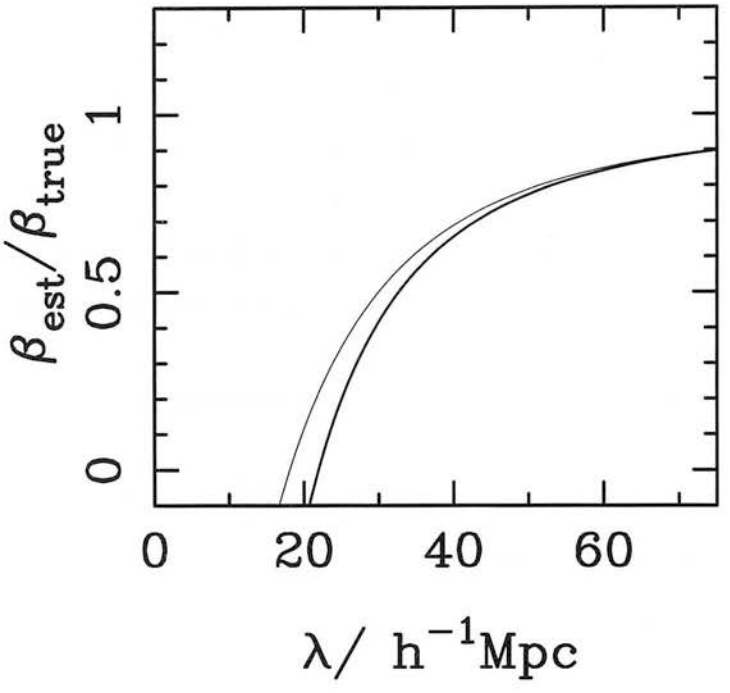

Figure 2.5: The ratio of estimated to true β plotted against wavelength (scale), with $\beta = 0.5$ and 'Fingers of God' models with $\sigma_p = 450\text{km s}^{-1}$. The thick curve is for the the Gaussian model, the thin curve for the exponential model.

If the non-linear 'Fingers of God' effect is included, it dramatically alters the redshift space power spectrum shape – see Fig 2.4. It can be seen that the linear power spectrum boost is reduced and actually reversed at small enough scales, with the contours appearing flattened along the line of sight. The two non-linear models appear to have rather different effects in this figure. However it must be remembered that medium and large scales are all compressed into the bottom left corner. The quantitative effect of non-linear distortions on the CFW94 result can be seen in Fig 2.5. As expected, the two models converge at large enough scales (good agreement where $k\sigma_p < 1$), see equation (2.27). However, it is clear that the β estimate suppression is significant out to very large scales ($\sim 60\,h^{-1}\,\text{Mpc}$). This curve is quite a good fit to the N-body simulation result for $\beta = 0.5$ in CFW94 (except perhaps at large scales) – see Fig 2.6. This shows that the model predicts the 'Fingers of God' effect quite well.

It is unclear how good the fit is, because the CFW94 results are noisy – there are two independent simulations which diverge quite significantly, particularly at large scales. However, some comparison can be made with the model results. The quoted velocity dispersion for this simulation is the higher value of $\sigma_p = 600\text{km s}^{-1}$. For this high value of dispersion, the model predicts a stronger suppression of the β estimate – see Fig 2.7. This curve is perhaps a better fit to the large scale end of the CFW94 result, although their curve is particularly noisy at that end. Results at the large-scale end have larger errors as there are fewer independent modes ($n_{\text{modes}} \propto k^3$ – this is analogous to the density of states in statistical mechanics).

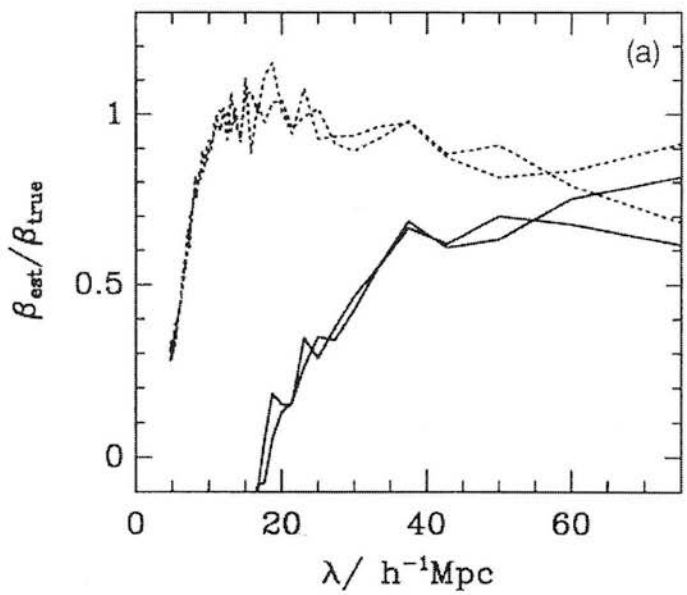

Figure 2.6: The β ratio curves from the CFW94 simulations with $\beta = 0.5$ and a measured $\sigma_p = 600\text{km s}^{-1}$. The dotted curves are for smoothed particle velocities and should be ignored here.

The reduced effect of small scale dispersions at some scales in the simulation may be related to the uncertainty in the meaning of σ_p. For example there may be small scale velocity correlations, which would reduce the degree to which power is 'smeared out'. The model assumption that small scale velocities are totally uncorrelated is only a crude approximation.

CFW94 also perform a simulation with $\Omega = 0.3$ and $b = 2$, *i.e.* $\beta = 0.24$, see

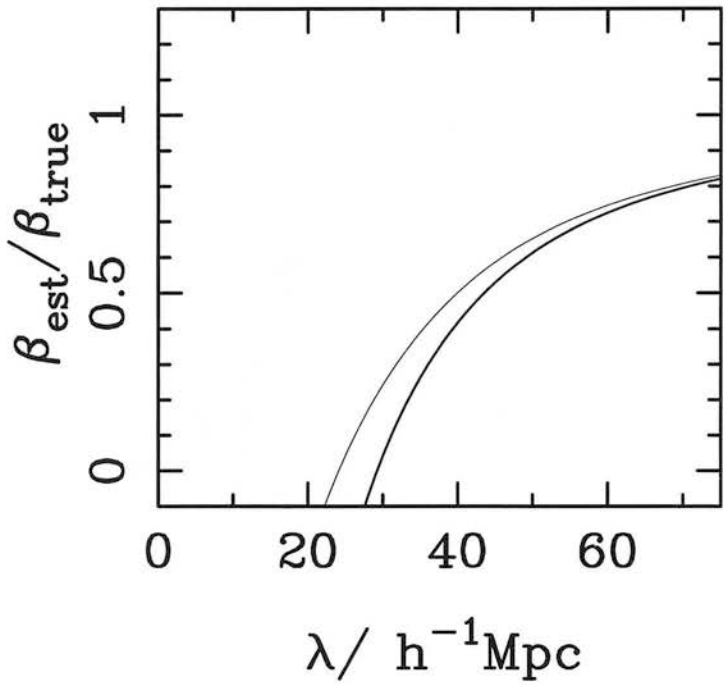

Figure 2.7: β ratio curve with $\beta = 0.5$ and $\sigma_p = 600\text{km s}^{-1}$.

Fig 2.8 This is also simulated using our model. CFW94 quotes the velocity dispersion here as $\sigma_p = 250\text{km s}^{-1}$. The theoretical curves for these parameters are given in Fig 2.9 Again it isn't a very good match except at large scales, probably because of velocity correlations. A better fit to the CFW94 $\beta = 0.24$ curve can be had with $\sigma_p = 160\text{km s}^{-1}$, which fits quite well, except at large scales.

The problem with the analysis of CFW94 is that it is based on the Kaiser approximations of linear theory and a distant observer. Linearity requires large wavelength modes, at least 80Mpc (as can be seen from Fig 2.5), and hence at least as large a survey. The distant observer approximation requires that the survey volume subtends a small angle on the sky *i.e.* that it is very much further from the observer than its scale. For accuracy this would require a survey at a huge distance, where current instruments cannot detect a good sample of galaxies. These 'Finger of God' models could be used to improve accuracy at non-linear scales, as was done by CFW95. Their simulations showed that the approximation seemed to work quite well, but further work by Hatton & Cole (1997) casts some doubt on this, and it may depend in detail on things such as

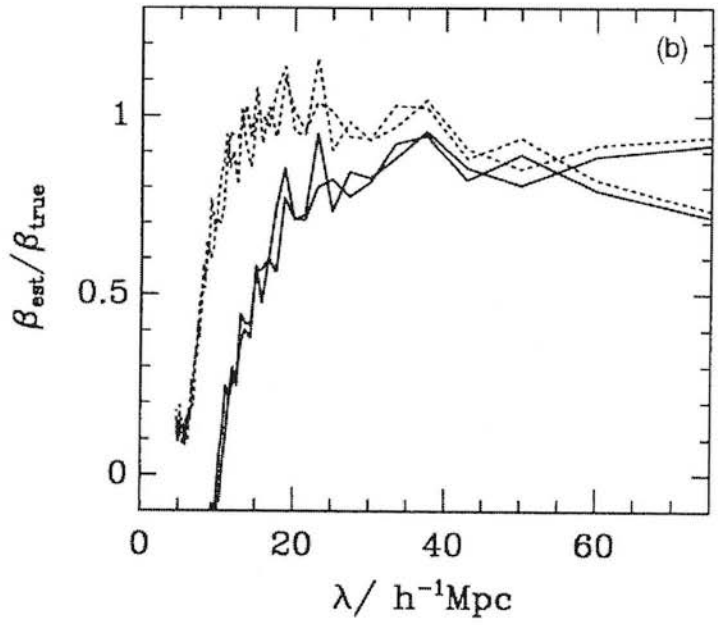

Figure 2.8: CFW94 β ratio N-body results with $\beta = 0.24$ and $\sigma_p = 250\text{km s}^{-1}$. Ignore the dotted curves.

bias.

The value σ_p parameter is uncertain, and very difficult to measure accurately – see Mo *et al.* (1993). In principle it is possible to fit a value of σ_p from data taken at several different values of λ. However, the models assume that these pairwise velocity fields are incoherent which is not entirely true, and this coherence may cause significant deviations from our simple models. It appears to be necessary to study scales larger than those allowed by the distant observer/linearity tradeoff which is inherent in this kind of anisotropic power spectrum analysis.

Clearly the ideal thing to do is analyse the largest modes in the survey, which will be most reliably linear and have the additional advantage that linear biasing should be a more reliable approximation. Thus it is better to use the whole survey at once and decompose the density field in spherical polar coordinates – see chapter 4 and chapter 5.

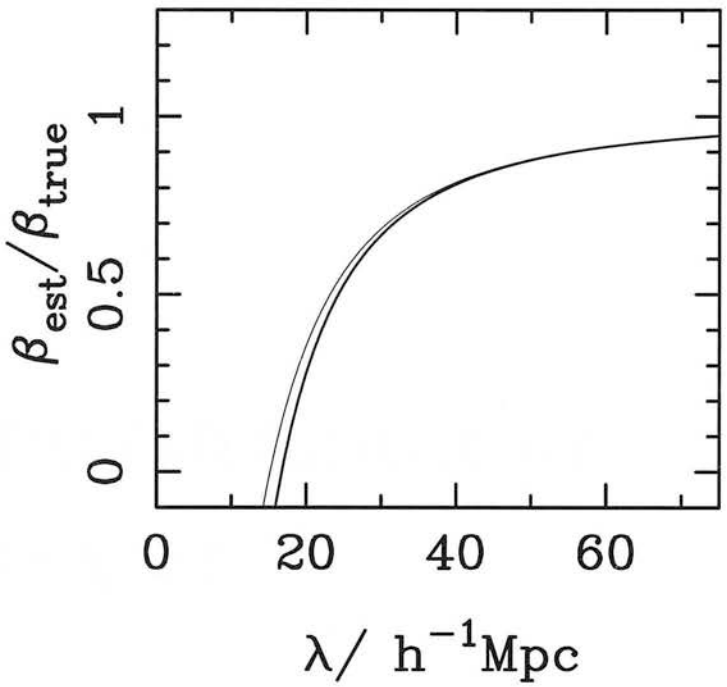

Figure 2.9: β ratio curve with $\beta = 0.24$ and $\sigma_{\rm p} = 250{\rm km\ s^{-1}}$.

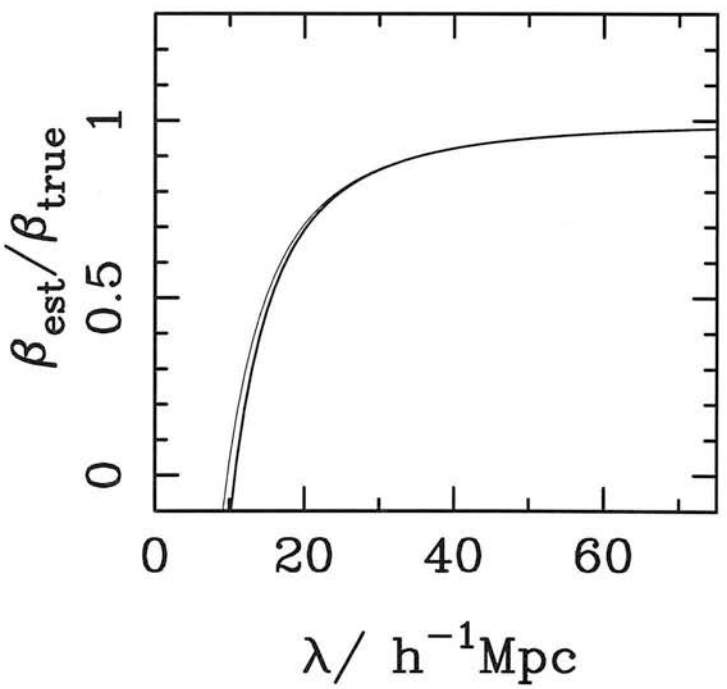

Figure 2.10: β ratio curve with $\beta = 0.24$ and $\sigma_{\rm p} = 160{\rm km\ s^{-1}}$.

Chapter 3

THE COSMOLOGICAL CONSTANT

This chapter is taken from the published paper Ballinger, Peacock & Heavens (1996).

3.1 Introduction

In recent years there has been a resurgence of interest in the cosmological constant Λ as a possible way of evading several cosmological problems (see Carroll, Press & Turner 1992 for a review). Although long popular with theorists, the Einstein-de Sitter $\Omega_{\rm m} = 1$ model (hereafter denoted EdS) now seems increasingly untenable owing to its short expansion time-scale. An open model with low matter density parameter $\Omega_{\rm m}$ and no cosmological constant allows a younger universe, but is more difficult to reconcile with inflation. Although some workers have suggested that bubble nucleation within inflation may be capable of yielding an open universe (e.g. Bucher, Goldhaber & Turok 1995), a more common alternative to the EdS model is to retain a flat universe, $k = 0$, through $\Omega \equiv \Omega_{\rm m} + \Omega_\Lambda = 1$ (where $\Omega_\Lambda = \Lambda/3H_0^2$). A high value of $\Omega_\Lambda \simeq 0.8$ would then be indicated by arguments for $\Omega_{\rm m} \simeq 0.2$ from cluster mass-to-light ratios and large-scale structure measurements (Efstathiou, Sutherland & Maddox 1990). In any

case, it must be remembered that there is no known reason for the vacuum density to vanish (Weinberg 1989), and so it is profoundly important for physics to test whether Λ is non-zero.

It has been suggested by Alcock & Paczyński (1979) that it may be possible to detect the presence of Λ by a geometric test, measuring the effect of deviations from the assumed EdS geometry on large-scale structure. The assumption of an incorrect geometry can lead to an effective squashing of space along the line of sight – causing an anisotropy in the inferred density field which could be detected from galaxy (or quasar) clustering statistics. This has the crucial advantage over other tests such as number counts of being independent of galaxy or quasar evolution. Phillipps (1994) considered the possibility of analysing a quasar survey using the orientation of pairs and claimed that the effect should be readily detectable. However, the only cosmological constant model discussed by Phillipps is de Sitter space: zero mass density and $\Omega_\Lambda = 1$. We show that if the matter density parameter is even modestly greater than zero then the geometric effect is much reduced and also does not continue to increase with redshift above $z \simeq 1$.

Redshift-space distortions caused by peculiar velocities of galaxies also lead to anisotropic structure. Large-scale infall squashes overdensities along the line of sight in redshift space, which can mimic the geometric squashing caused by Λ. In the linear regime this distortion has a simple effect in Fourier space in the 'distant observer' approximation (see Kaiser 1987) and is characterized by the parameter $\beta = \Omega_{\rm m}^{0.6}/b$ (where b is the bias parameter). For parameters of interest, redshift distortion is not negligible in comparison with the geometric distortion. The situation is complicated further on small scales, where virialized clusters appear elongated along the line of sight – the so-called 'fingers of God'.

We therefore consider in some detail the clustering anisotropies which arise in the presence of all three effects: geometric flattening, β-distortion and fingers of God. We use a power-spectrum analysis, since the modelling of redshift-space distortions is simpler in Fourier space than it is in real space. Although the Kaiser and squashing effects are similar, the functional forms of the anisotropies in the power spectrum differ,

and Λ and β can be distinguished in principle. Sections 3.2 and 3.3 plus Appendix B give the basics of the effect. In Section 3.4, we present a maximum-likelihood technique for measuring both Λ and β from these anisotropies in Fourier space. Finally, in Section 3.5, we give some assessment of the likely practical constraints on Λ that may be expected from next-generation redshift surveys, such as the Anglo-Australian 2-degree field survey and the Sloan Digital Sky Survey.

3.2 Cosmological Models and Clustering Anisotropy

To measure the cosmological constant, we exploit the fact that we cannot measure comoving distances directly, but use redshift as a distance indicator. An object that is spherical in comoving real space $\boldsymbol{r}$ will only appear spherical to an observer (who measures redshifts) if the correct geometry is assumed – *i.e.* the correct $r(z)$ relation is used.

We write the Robertson-Walker metric as

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - R(t)^2\left[\mathrm{d}r^2 + S_k^2(r)\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2\right)\right]. \tag{3.1}$$

Here $R_0 r$ is the comoving geodesic distance, $R(t)$ is the cosmic scale factor ($= R_0$ now) and

$$S_k(r) = \begin{cases} \sin r & k = 1 \quad \text{closed} \\ r & k = 0 \quad \text{flat} \\ \sinh r & k = -1 \quad \text{open.} \end{cases} \tag{3.2}$$

The square of the comoving distance between two objects at $z \pm \Delta z/2$ separated by $\Delta\theta$ is

$$\Delta\ell^2 = R_0^2 S_k(r)^2 \Delta\theta^2 + R_0^2 \left(\frac{\mathrm{d}r}{\mathrm{d}z}\right)^2 \Delta z^2. \tag{3.3}$$

For a universe with matter contribution Ω_m and cosmological constant contribution Ω_Λ:

$$\begin{aligned} R_0\frac{\mathrm{d}r}{\mathrm{d}z}(z) &\equiv A_\mathrm{t}(z) \\ &= \frac{c}{H_0}\frac{1}{\sqrt{(1-\Omega)(1+z)^2 + \Omega_\Lambda + \Omega_\mathrm{m}(1+z)^3}}, \end{aligned} \tag{3.4}$$

$$R_0 S_k(r) \equiv B_{\rm t}(z) = \frac{c}{H_0}|1-\Omega|^{-1/2} \times S_k\left[\int_0^z \frac{|1-\Omega|^{1/2}\,{\rm d}z'}{\sqrt{(1-\Omega)(1+z')^2+\Omega_\Lambda+\Omega_{\rm m}(1+z')^3}}\right]. \tag{3.5}$$

For flat models, favoured by inflation, $\Omega = \Omega_{\rm m} + \Omega_\Lambda = 1$, so

$$R_0 \frac{{\rm d}r}{{\rm d}z}(z) = \frac{c}{H_0}\frac{1}{\sqrt{(1-\Omega_{\rm m})+\Omega_{\rm m}(1+z)^3}}, \tag{3.6}$$

$$R_0 r(z) = \frac{c}{H_0}\int_0^z \frac{{\rm d}z'}{\sqrt{(1-\Omega_{\rm m})+\Omega_{\rm m}(1+z')^3}}. \tag{3.7}$$

Even with this simplification, the integral for $R_0 r(z)$ must be carried out numerically unless $\Omega_\Lambda = 0$ or $\Omega_{\rm m} = 0$.

There are three coordinate systems to consider here. We denote by the subscript t the true geometry, where z represents the redshift from pure Hubble expansion at the coordinate distance r. We denote by subscript s the coordinate system where the geometry is correct, but the redshift is used as a distance indicator. This incorporates redshift distortion from peculiar velocities. Finally, we denote by the subscript a the assumed geometry. This will always be an EdS model in the present chapter, and quoted flattening factors are given relative to this model, for which we have

$$R_0 \frac{{\rm d}r}{{\rm d}z}(z) = \frac{c}{H_0}\frac{1}{(1+z)^{3/2}} \equiv A_{\rm a}(z), \tag{3.8}$$

$$R_0 r(z) = 2\frac{c}{H_0}\left(1 - \frac{1}{\sqrt{1+z}}\right) \equiv B_{\rm a}(z). \tag{3.9}$$

Taking EdS as default, we define geometric squashing factors:

$$f_{\parallel}(z) \equiv \frac{A_{\rm t}}{A_{\rm a}}, \tag{3.10}$$

$$f_{\perp}(z) \equiv \frac{B_{\rm t}}{B_{\rm a}}, \tag{3.11}$$

$$F(z) \equiv \frac{f_{\parallel}}{f_{\perp}} = 1 + \frac{1}{4}(1 - \Omega_{\rm m} + 2\Omega_\Lambda)\,z + O(z^2). \tag{3.12}$$

$f_\perp, f_\parallel$ and F are all unity for EdS. The flattening factor F is defined so that $F > 1$ means that objects would appear flattened along the line of sight in the assumed geometry. The first-order redshift dependence (obtained with the aid of MATHEMATICA) tells us that what is really measured via the flattening is not Ω_Λ but $\Omega_\Lambda - \Omega_m/2$. This is inevitable, since a general argument (e.g. section 14.6 of Weinberg 1972) shows that the lowest-order corrections to the distance-redshift relation depend only on $q_0 = \Omega_m/2 - \Omega_\Lambda$. The accuracy with which one can measure Ω_Λ from low-redshift data is then limited by (amongst other things) one's knowledge of Ω_m, which is currently rather poorly known. In principle, measuring the evolution at several redshifts out to $z > 1$ would allow both Ω_m and Ω_Λ to be determined separately, but this will be harder. It is also interesting to compare the Ω_m dependence in the case of open and flat models: $F \simeq 1 + (1 - \Omega_m)z/4$ with $\Omega_\Lambda = 0$ but the three-times stronger $F \simeq 1 + 3(1 - \Omega_m)z/4$ for $k = 0$.

It can be seen from Figs 3.1(a) and (c) that the squashing effect of Λ (characterized by F) increases rapidly with redshift and then saturates at redshift of $z \simeq 1$ for all but pure de Sitter space ($\Omega_m = 0$). The fact that F does not increase without limit if $\Omega_m \neq 0$ is easily understood. At high redshifts matter dominates over vacuum energy because of their equations of state, and

$$\frac{\Omega_m(z)}{\Omega_\Lambda(z)} \propto (1+z)^3. \tag{3.13}$$

Hence, at high z, the cosmological constant becomes insignificant. The saturation is quite severe: $F \leq 1.3$ for $\Omega_m \geq 0.2$ (the limit of most current models). Compare this with the de Sitter model which reaches $F = 2.2$ at $z = 2$. There is thus unlikely to be any gain by going to very high redshifts: a galaxy survey with a large number of galaxies may well be preferable to a deeper but much smaller quasar survey. The calculations can be repeated for open models with no cosmological constant, but the effects are much smaller (see Fig. 3.1b). Hence a large value of F indicates a cosmological constant, and would be a valuable robust measure of Λ, if it were not subject to confusion with other effects.

To utilize the squashing as a diagnostic for the geometry, it is necessary to have

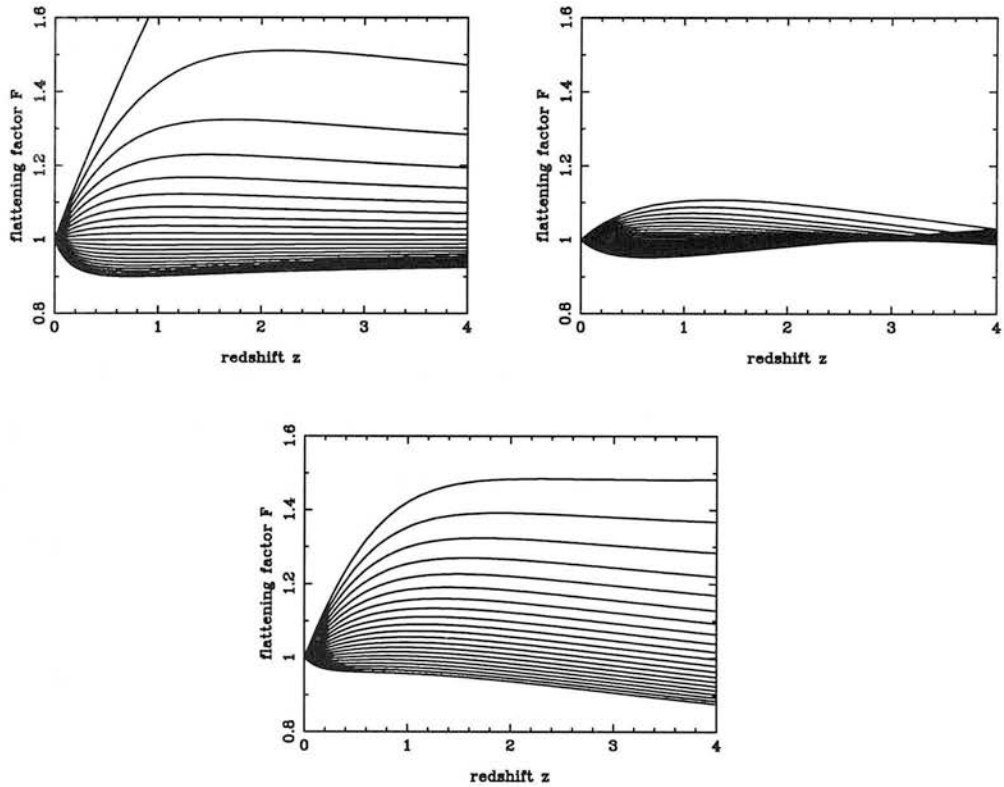

Figure 3.1: (a) Flattening factor F as a function of redshift for flat models ($\Omega_{\rm m} + \Omega_\Lambda = 1$). The curves range from $\Omega_\Lambda = 1$ (top curve) to $\Omega_\Lambda = -1$ in steps of 0.1. (b) F against redshift for models with no cosmological constant. $\Omega_{\rm m}$ varies from 0 (top) to 2, again in steps of 0.1. (c) F against redshift for models with a fixed mass density $\Omega_{\rm m} = 0.2$ and a cosmological constant. Ω_Λ varies as for (a).

something of known shape. A convenient such object is the power spectrum of galaxy (or quasar) clustering, whose contours in **k**-space must be spherical if the cosmological principle holds (although a survey that covered a significant range of redshifts would require power to be measured separately in a number of redshift bins to avoid anisotropy introduced by evolution). Denoting the real comoving wavenumber by **k**, and splitting it into a component $k_{||}$ along the line of sight and a component $k_\perp$ perpendicular to it, the true comoving real-space power spectrum will be a function only of $|k|$:

$$P_{\rm t}(k_{||}, \mathbf{k}_\perp) = P_{\rm t}(k). \tag{3.14}$$

If we assume the wrong geometry, we measure power at the wrong wavelengths. Ignoring redshift distortions for the time being:

$$P_{\rm a}(k_{||}, {\rm k}_\perp) = \frac{1}{f_\perp^{3+n} F} P_{\rm t}(k) \left[1 + \mu_{\rm a}^2 \left(\frac{1}{F^2} - 1\right)\right]^{\frac{n}{2}}, \tag{3.15}$$

where n is the local spectral index of the power spectrum and μ_a = cosine of wavevector to line of sight. This expression is derived in Appendix B, but is immediately reasonable: $P(\mu_{\rm a} = 0)/P(\mu_{\rm a} = 1) = F^n$, as expected for squashing by a factor F. This equation implies that there is no sensitivity to geometry for an $n = 0$ spectrum; however, this ceases to be true in the presence of peculiar velocities, which also give rise to power-spectrum anisotropies.

3.3 Redshift Distortions

3.3.1 Linear modes

In models where structure forms via gravitational instability, density fluctuations inevitably induce peculiar velocities, which affect the mapping to redshift space where the radial coordinate is a total velocity. In linear theory, density waves are boosted along the line of sight in redshift space (Kaiser 1987). We assume that the survey subtends a small angle in the sky and lines of sight can be treated as parallel, so that the simple Kaiser formula remains valid (see Heavens & Taylor 1995 and Zaroubi & Hoffman 1996 for more general analyses) – *i.e.* even with the correct assumed geometry, there is still anisotropy in redshift space:

$$P_{\rm s}(k, \mu) = P_{\rm t}(k)[1 + \beta\mu^2]^2 \tag{3.16}$$

where $P_{\rm t}$ is the isotropic real-space power spectrum and $P_{\rm s}$ is the power spectrum in the correct geometry.

This boosts the power spectrum along the line of sight as does the Λ effect, but it is

clear from formulae (3.15) and (3.16) that the μ dependence is different. Gravitational instability generates a characteristic ratio between the μ^2 and μ^4 components, which differs in general from that resulting from geometrical distortion. Hence the two effects are distinguishable in principle, given data of sufficiently high signal-to-noise ratio (see Fig. 3.2).

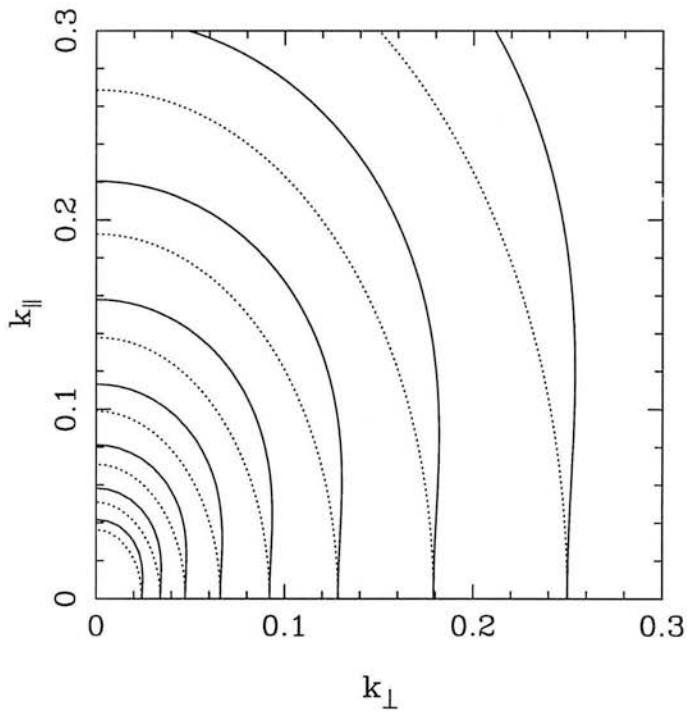

Figure 3.2: Contours of the power spectrum in the $k_{||}$ and $k_\perp$ plane, assuming a power-law index $n = -1.5$ for the true power spectrum. The contour interval is $\Delta \ln P = 1/2$. The models are linear (Kaiser) redshift distortion only with $\beta = 0.5$ (full contours) and geometric squashing effect only, with $F = 1.5$ (dotted contours).

3.3.2 Fingers of God

Additionally, we consider the non-linear finger of God effect caused by velocity dispersions in virialized clusters. A simple model for the effects of such velocities on the power spectrum was introduced by Peacock (1992; see also Peacock & Dodds 1994). This consists of multiplying the linear-theory Kaiser distortion by a term which treats

the radial distortion as a convolution with an incoherent velocity component, leading to damping of power at large values of $k_{||}$:

$$P_s(k,\mu) = P_t(k)[1+\beta\mu^2]^2\ D[k\mu\sigma_p]. \tag{3.17}$$

If the dispersion is taken to be Gaussian, the damping term is

$$D[k\mu\sigma_p] = \exp[-k^2\mu^2\sigma_p^2/2]. \tag{3.18}$$

Here σ_p is the line-of-sight *pairwise* dispersion in relative galaxy velocities caused by the incoherent dispersion ($\sigma_p = \sqrt{2}\sigma_v$). This is usually quoted in velocity units, but is implicitly divided by 100 to obtain a scale-length in h^{-1}Mpc when used in power spectra[1]. Note that it is not obvious that this parameter corresponds exactly to the pairwise velocity dispersion measured using other methods, although it should be close. In reality, the pairwise velocity distribution is better modelled by an exponential (Davis & Peebles 1983), which leads to a Lorentzian factor in Fourier space:

$$D[k\mu\sigma_p] = \frac{1}{1+(k\mu\sigma_p)^2/2}. \tag{3.19}$$

In practice, there is very little difference between these models until the damping becomes a factor $\gtrsim 2$ – see Fig. 3.3. According to a comparison with N-body simulations by Cole, Fisher & Weinberg (1994 and 1995) (hereafter CFW94 and CFW95), this simplified analytical model appears to work well up to $k\sigma_p$ of order unity. We note that CFW95 in fact use a damping factor of a Lorentzian squared, which is not equivalent to an exponential pairwise distribution (it actually corresponds to an exponential one-particle distribution); however, this makes very little difference to the answers.

3.3.3 The general case

Combining the effects of geometry, linear peculiar velocities and fingers of God, it is shown in Appendix B that the overall effect is

$$\begin{aligned} P_a(k,\mu) &= \frac{1}{f_\perp^{3+n}F}P_t(k)\left[1+\mu^2\left(\frac{1}{F^2}-1\right)\right]^{\frac{n-4}{2}} \\ &\times \left[1+\mu^2\left(\frac{\beta+1}{F^2}-1\right)\right]^2 D[k\mu\sigma_p'], \end{aligned} \tag{3.20}$$

[1]Hubble constant $H_0 = 100\,h$ km s^{-1}Mpc^{-1}.

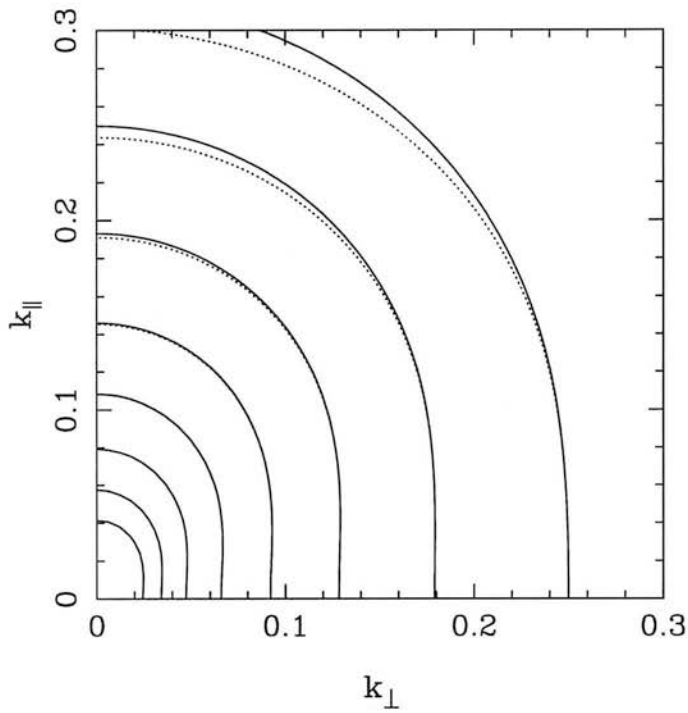

Figure 3.3: Contours for both linear and non-linear redshift distortions with $\beta = 0.5$, and $\sigma_p = 350$ km s^{-1} (again for $n = -1.5$). The solid contours are for exponential small scale velocity distribution, the dotted Gaussian. The contour interval is $\Delta \ln P = 1/2$. The stretching of contours of the correlation function along the line of sight leads to a reduction in line of sight power on small scales (large k) – see also Cole, Fisher & Weinberg (1994).

where $\sigma_p' \equiv \sigma_p / f_{||}$, and we have modified μ to account for the assumed geometry. If $\beta = 0$, and F is close to unity, then the anisotropy resembles a pure redshift-space distortion, with an apparent value of β:

$$\beta_a \simeq -\frac{n}{2}(F - 1). \tag{3.21}$$

If the true β is small but non-zero, then the true distortion adds linearly to the geometrical signature; for large β, things are more complex, since F also modifies the effect of β. Since the effective $n \simeq -1$ on the scales of interest where β can be measured, we see that any experiment which hopes to detect Λ through $F \simeq 1.3$ will need to be able to measure β_a to a precision of rather better than ± 0.1. Fig. 3.4 shows that the differences

between the F and β effects are subtle; a large survey with good statistical signal-to-noise ratio will be needed if the two effects are to be measured separately. Furthermore, because the effects of interest are small, careful control of systematics will be required. For example, one will not wish to assume the precise correctness of the model (3.17) for the redshift-space distortions, and empirical relations from numerical simulations will be more appropriate. However, for the present the simple model suffices to indicate how hard we will have to work. It also illustrates how well one needs to know the intrinsic shape of the objects whose flattening is being exploited. The scatter in the shapes of cosmic objects such as voids (cf. Ryden 1995) will preclude their use for this purpose.

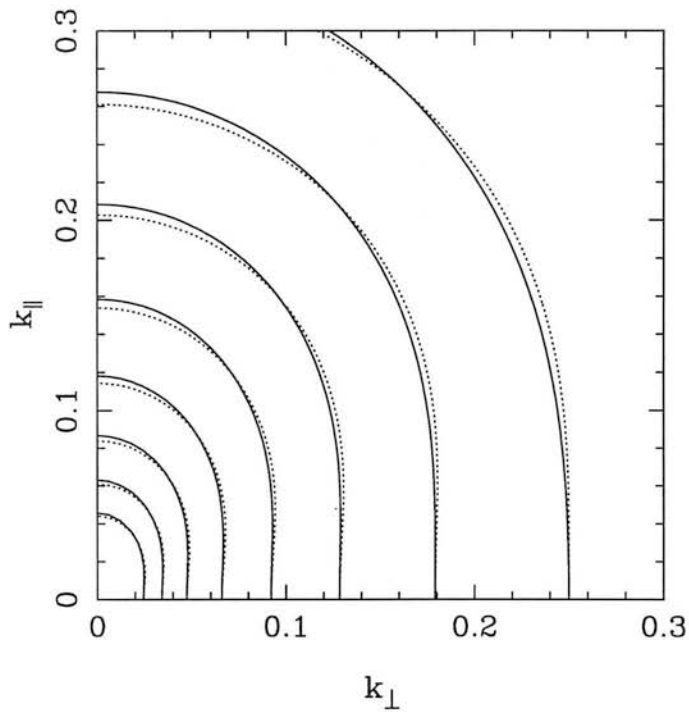

Figure 3.4: Solid contours for a model with $F = 1.1$, $\beta = 0.5$ and $\sigma_p = 350$ km s^{-1}. Dotted contours are for the best fit to this model with redshift distortions only ($F = 1$).

3.3.4 Geometry from evolution of β

Although we have shown that F and β are degenerate to some extent at a single redshift, they are expected to evolve differently. This in itself does not help to detect Λ, because the evolution of bias is unknown:

$$\beta(z) = \frac{\Omega_{\rm m}^{0.6}(z)}{b(z)}. \tag{3.22}$$

Nusser & Davis (1994) have argued that $b(z)$ is calculable provided galaxy numbers are preserved; however, we feel it is better not to assume that merging is negligible and to treat the evolution of bias empirically. This argues that we must seek a signature of Λ which takes the empirical $\beta(z)$ without seeking to assume what fraction of the evolution is due to changes in $\Omega_{\rm m}$ with redshift. The following sections of this chapter are devoted to seeing what can be done by looking at the detailed μ dependence of the clustering anisotropies, in order to break the lowest-order degeneracy.

However, further information on $b(z)$ can be obtained from the evolution of the amplitude of the power spectrum with redshift, as discussed by Matsubara & Suto (1996). In the linear regime, the true galaxy power spectrum evolves as

$$P_{\rm t}(z) = b^2(z) P_{\rm mass}(z) = \frac{b^2(z)}{b^2(0)} \mathcal{D}^2(z)\, P_{\rm t}(0), \tag{3.23}$$

where $\mathcal{D}(z)$ is the linear-theory growth law for density perturbations. Unfortunately, this evolution is not directly measurable: even looking at real-space clustering, there are geometrical factors which alter the inferred clustering (cf. the $\mu = 0$ limit of equation 3.20); what is observable is thus the apparent growth factor

$$\mathcal{D}_{\rm a}(z) = \frac{b(z)}{b(0)} \frac{\mathcal{D}(z)}{[f_\perp^{3+n}(z) F(z)]^{1/2}}. \tag{3.24}$$

The other observable is the apparent redshift-space distortion as a function of redshift, which we choose to describe by an apparent value of β. If the true β is of order unity, then the geometrical and velocity distortions interfere in a complex manner. Suppose we deduce an apparent β by fitting $P(\mu)/P(\mu = 0)$ by a velocity-only distortion; to first order in $(F - 1)$, this gives an apparent value of β:

$$\beta_{\rm a}(\mu) = \beta - \left[2\beta(1 - \mu^2) + \frac{n}{2}(1 + \beta\mu^2) \right] (F - 1). \tag{3.25}$$

There are thus many possible choices of β_a; the simplest procedure is probably to average these choices over μ, yielding the redshift-dependent observable

$$\beta_a(z) = \beta(z) - \frac{8\beta(z) + [3 + \beta(z)]n}{6} (F(z) - 1). \tag{3.26}$$

For a given hypothetical $(\Omega_m, \Omega_\Lambda)$ pair, the equations for $\beta_a(z)$ and $\mathcal{D}_a(z)$ both yield $b(z)/b(0)$; requiring these curves to match at all z can in principle allow both density parameters to be determined.

It is interesting to explore this possibility in a little more detail by looking at the first-order change with redshift of these relations. The initial changes with redshift of the observables are respectively

$$\frac{d\mathcal{D}_a}{dz} = -\Omega_m^{0.6} - \frac{n+4}{8}[1 - \Omega_m + 2\Omega_\Lambda] + \frac{d\ln b}{dz} \tag{3.27}$$

and

$$\frac{d\beta_a}{dz} = \left[\frac{4\beta}{15} - \frac{n}{24}(3+\beta)\right](1 - \Omega_m + 2\Omega_\Lambda) - \beta\frac{d\ln b}{dz}. \tag{3.28}$$

The second of these depends just on q_0, but the first involves a different combination of Ω_m and Ω_Λ. If we eliminate the unknown bias evolution $d\ln b/dz$ from the above equations, we get

$$\Omega_m^{0.6} = -\frac{d\mathcal{D}_a}{dz} - \frac{d\ln\beta_a}{dz} - \left[\frac{7}{30} + \frac{n}{6} + \frac{n}{8\beta_a}\right](1 - 2q_0). \tag{3.29}$$

For plausibly realistic values $n \simeq -1$, $\beta \simeq 0.5$, the coefficient of $(1 - 2q_0)$ is very small, giving a diagnostic for Ω_m alone. Studies of clustering evolution and anisotropy out to $z \simeq 0.5$ may thus principally determine Ω_m, rather than Ω_Λ. In order to pin down the cosmological constant, it will be necessary (a) to work at $z \gtrsim 1$; or (b) to work at large scales where $n \simeq 0$; or (c) to look in more detail at the angular dependence of the anisotropies. We now return to the last of these options.

3.4 Maximum-Likelihood Fitting

Given the model (3.20) for the combined geometrical plus redshift-space distortions, the standard way of proceeding would be to perform a maximum-likelihood fit on a

dataset to obtain the required parameters β, F, $\sigma_{\rm p}' = \sigma_{\rm p}/f_{||}$ and a parametrized form for $P_{\rm t}(k)$ simultaneously. Smaller uncertainties in the most interesting parameters such as β and F may be obtained if we use a priori knowledge of e.g. the shape of $P_{\rm t}$, but, in order to make a robust test for Λ, it is safer to obtain all the information we require from the same survey.

To apply maximum-likelihood methods, we need to know the probability distribution for the power. The central limit theorem implies that the observed power averaged over many independent modes will have a Gaussian distribution (even though single modes will be Rayleigh distributed for a Gaussian random field – see Feldman, Kaiser & Peacock 1994; hereafter). Similarly, $\ln P$ would be expected to have a Gaussian distribution with variance $\sigma^2 = 1/N$ where N is the number of independent modes averaged over:

$$f(\ln P_{\rm a})\,{\rm d}\ln P_{\rm a} = \left(\frac{N}{2\pi}\right)^{\frac{1}{2}} \times \exp\left[-\frac{N}{2}(\ln P_{\rm a} - \ln P_{\rm a}')^2\right]{\rm d}\ln P_{\rm a}, \tag{3.30}$$

where $P_{\rm a}$ is the observed power as a function of k and μ; $P_{\rm a}'[k,\mu;\Lambda,\beta,\sigma_{\rm p}',P_{\rm t}(k)]$ is the corresponding mean power of the model. In both cases, the powers are the *total* power, including shot noise. The likelihood estimator is then

$$\ln\mathcal{L}[\Lambda,\beta,\sigma_{\rm p}',P_{\rm t}(k)] = {\rm const.} - \frac{1}{2}\sum_{\rm modes}(\ln P_{\rm a} - \ln P_{\rm a}')^2; \tag{3.31}$$

in the continuum approximation, this becomes

$$\ln\mathcal{L} = {\rm const.} - \rho\int_0^{k_{\rm max}}\pi k^2\,{\rm d}k\int_0^1(\ln P - \ln P')^2\,{\rm d}\mu, \tag{3.32}$$

where ρ is the density of states in **k**-space.

For a non-uniform selection function, the effective density of states is (Appendix C)

$$\rho_{\rm parameterised} = \frac{[\int w\,\bar{n}\,{\rm d}^3r]^2}{(2\pi)^3\int w^2\bar{n}^2\,{\rm d}^3r}, \tag{3.33}$$

where $\bar{n}(r)$ is the mean density of objects, and $w(r)$ is the optimal weighting function of Feldman et al. (1994) for measuring power: $w = (1+\bar{n}P)^{-1}$. The integral is restricted to a maximum value of $k = k_{\rm max}$ to prevent excessive non-linearity, in particular to avoid wavenumbers where the ansatz for the fingers of God may not be a good approximation. CFW95 showed that this limit corresponds to $k\sigma_{\rm p}' \lesssim 1.6$ (10 per cent error in the derived β). The precise maximum wavenumber allowed by this criterion depends on the assumed $\sigma_{\rm p}$. Taking a conservative view of this, we will generally set $k_{\rm max} = 0.35\,h\,{\rm Mpc}^{-1}$ as a safe limit within which F is not significantly affected by systematics in the inferred β. The sum/integral is over a hemisphere in k-space. The modes in the other hemisphere are identical: $P(\mathbf{k}) = P(-\mathbf{k})$ because the density field is real.

The procedure in practice is to maximize $\mathcal{L}$ to find the best-fitting parameters Λ, β, $\sigma_{\rm p}'$ and $P_{\rm t}(k)$. We assume a shape for $P_{\rm t}(k)$ and fit only for the amplitude, although the shape could be parametrized for a more extensive analysis. In this chapter we are interested principally in the errors on the parameters, to work out whether any given survey will be able to distinguish between cosmological models of interest.

3.4.1 Covariances and correlation matrix

We now wish to study the form of the likelihood contours that describe the joint distribution of the interesting parameters in this problem. This question can be studied in the absence of data because of the simple form of the likelihood function, which is a sum of squared differences in $\ln P$, multiplied by the density of states in k-space. For a given observed power, the likelihood thus scales simply with the volume of the survey; the likelihood contours are of a fixed form, which are merely re-labelled as the number of modes increases. The covariance matrix C_{ij} for the parameters a_i is thus of a universal shape which can be scaled once the absolute value of the error on one parameter is known (at least for a fixed ratio of true power to shot power).

Even the effect of shot noise is rather weak. It only enters at all because the error in power scales with the power: $\sigma(P)$ thus changes with μ in the presence of redshift-space anisotropies, provided shot noise is negligible. Conversely, if shot noise

dominates, $\sigma(P)$ is a constant. In one limit, therefore, the likelihood measures the sum of $[\Delta \ln P]^2$, whereas in the other it is $[\Delta P]^2$. In practice, the difference seems to be small, as shown in Figs 3.5 and 3.6. The optimally weighted QDOT dataset has $P_{\rm shot} = P_{\rm true}$ at $k = 0.05\ h\,{\rm Mpc}^{-1}$ and $P_{\rm shot} = 10P_{\rm true}$ at $k = 0.2\ h\,{\rm Mpc}^{-1}$. Future large-scale surveys are likely to have negligible shot noise on all quasi-linear scales.

To produce likelihood contours requires some data for the 'observed' power. In what follows, we simply take this to be the expected power in the model under consideration. A more exact procedure would be to produce a realization of the model, with the power for each mode exponentially distributed about the expectation value. The maximum-likelihood value of the parameters a_i obtained by fitting to such a fictitious dataset would then differ from the input ones, within the 'error bars' produced by the likelihood analysis. However, these errors are the same whether we input the expectation spectrum as data, or use a realization:

$$\begin{aligned} \Delta \ln \mathcal{L} &= -\frac{1}{2}\sum (\ln P_{\rm obs} - \ln P_{\rm mod})^2 \\ \Rightarrow C_{ij}^{-1} &= -\left.\frac{\partial^2 \ln \mathcal{L}}{\partial a_i \partial a_j}\right|_{\rm ML} \simeq \sum \frac{\partial \ln P_{\rm mod}}{\partial a_i}\frac{\partial \ln P_{\rm mod}}{\partial a_j}, \end{aligned} \tag{3.34}$$

independent of the data. Note that this expression is only approximately true, since it assumes that the term $\sum(\ln P_{\rm obs} - \ln P_{\rm mod})\partial^2 P_{\rm mod}/\partial a_i \partial a_j$ is negligible. The maximum-likelihood solution ensures that a somewhat different weighting of the data is zero: $\sum(\ln P_{\rm obs} - \ln P_{\rm mod})\partial P_{\rm mod}/\partial a_i = 0$. In what follows in this section, we neglect the additional term; the numerical results presented do not make this approximation. Using the expectation power spectrum as data therefore gives approximately the right error contours on the parameters, but centred on the input data. The constraints that can be set from a given real dataset would then depend on the actual maximum-likelihood values, which we would expect to be displaced from the true value by about the 'one-sigma' error calculated here.

Because the model power spectrum studied here is of a rather simple form, the covariance matrix can be obtained analytically in many cases, at least with the aid of a symbolic manipulator such as MATHEMATICA. The general expressions are sufficiently messy that they are not worth reproducing, but there are a few exact results that are simple enough to be useful. First recall that the power to be fitted is the sum of a

true power $P_{\rm true}$ plus shot power $P_{\rm shot}$; one might therefore approximate the covariance matrix as follows, simply interpolating between the scaling of the results for zero shot noise and for dominant shot noise:

$$C_{ij} = \frac{2}{N}\left(C_{ij}^{\rm A} + \frac{P_{\rm shot}}{P_{\rm true}} C_{ij}^{\rm B}\right), \tag{3.35}$$

where N is the total number of modes being considered (over a complete sphere in k-space). The covariance matrices for the shot-free and shot-dominated cases ($\mathbf{C}^{\rm A}$ and $\mathbf{C}^{\rm B}$ respectively) depend on the number of parameters being fitted for. Perhaps the most optimistic case would seem to be the two-parameter model where only β and F are allowed to vary, assuming that both the damping $\sigma_{\rm p}$ and the amplitude of the power spectrum are known. The latter assumption is probably unreasonable, since the main manifestation of a non-zero β or non-unity F is to change the mean level of power; the real signature we are interested in is the power anisotropy. Within the likelihood methodology, this demands that we allow the amplitude of the spectrum to float as a third parameter which is integrated over when we consider the distribution of the interesting parameters β and F.

For this model, the covariance matrices in the shot-free and shot-dominated limits can be obtained analytically in the limit of small β, where both have the same form:

$$C_{\beta\beta}^{A} = C_{\beta\beta}^{B} = \frac{11025\, n^2}{256\,(-4+n)^2}\,\beta^{-2}. \tag{3.36}$$

As we saw earlier when considering the quadrupole term, in this limit

$$C_{FF} = \frac{4}{n^2} C_{\beta\beta}, \tag{3.37}$$

so that β can be obtained with greater accuracy than F if $n \simeq -1$. If only β and F had been allowed to vary, both results would have been a factor 16/25 smaller – reflecting the spurious accuracy that arises if the effects on the normalization of the power spectrum are ignored. As β increases, $\beta^2\mathbf{C}$ changes approximately linearly with β, and it will suffice to quote the results at $\beta = 1$. For the shot-free case:

$$C_{\beta\beta}^{A}(\beta = 1) \simeq 22.4 - 16(n + 1.5), \tag{3.38}$$

$$C_{FF}^{A}(\beta = 1) \simeq 12.2 + 5(n + 1.5). \tag{3.39}$$

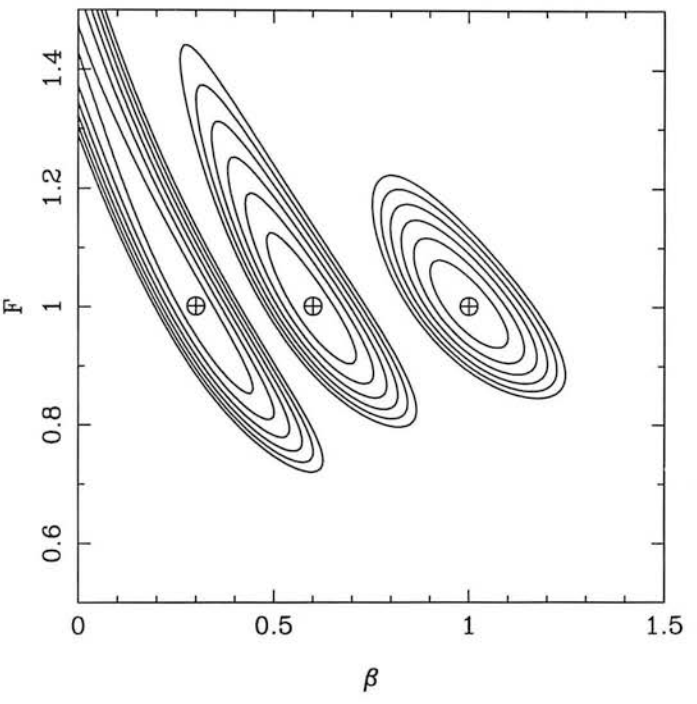

Figure 3.5: Expected contours of likelihood in the $F - \beta$ plane, for the case of a survey with negligible shot noise, for three values of β. The scaling is set so that $\sigma(\beta) = 0.1$ for $\beta = 1$. At each (β, F) point, the amplitude distribution has been integrated over. The contour interval is $\Delta \ln \mathcal{L} = 1/2$. The position of the true values of the parameters is marked by the cross; contours are shown for $\beta = 0.3$, 0.6 and 1.0.

For the shot-dominated case, the numbers are smaller:

$$C^B_{\beta\beta}(\beta = 1) \simeq 10.8 - 6(n + 1.5), \tag{3.40}$$

$$C^B_{FF}(\beta = 1) \simeq 2.7 + (n + 1.5). \tag{3.41}$$

We see that, although $C_{\beta\beta}$ is smaller than C_{FF} when β is small, $\beta^2 C_{\beta\beta}$ increases quite significantly with β, whereas $\beta^2 C_{FF}$ increases slowly or even declines. If β is large, it is therefore possible to measure F more precisely than β.

The correlations between the parameters are described by the correlation matrix

$$r_{ij} = C_{ij}/\sqrt{C_{ii}C_{jj}}, \tag{3.42}$$

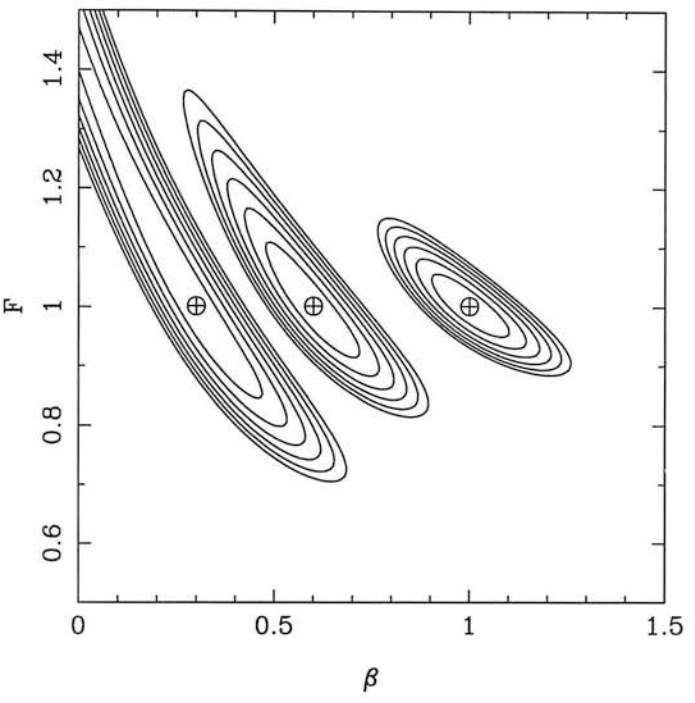

Figure 3.6: Expected contours of likelihood in the $F - \beta$ plane, as for Fig. 3.5 except now for the case of a survey dominated by shot noise.

where $C_{ij} = \langle \delta a_i \, \delta a_j \rangle$, and no summation is implied in C_{ii}. At $\beta = 0$, β and F are perfectly anticorrelated, with $r_{\beta F}$ increasing as β increases. In the shot-free case, $r_{\beta F}(\beta = 1) \simeq -0.75 + 0.4[1.5 + n]$, whereas $r_{\beta F}(\beta = 1) \simeq -0.85 + 0.2[1.5 + n]$ when shot noise dominates. This increase of the anti-correlation between F and β for lower values of β can be understood as follows: if we expand equation (3.20) in powers of μ, for small β and $\epsilon \equiv F - 1$, the coefficient of μ^2 is linear in β and ϵ, but gives degenerate information on these parameters. To separate β and ϵ requires at least the μ^4 term, which is of second-order smallness in β and ϵ.

These results all apply to long wavelengths, where the effects of the damping term are negligible. We give in Appendix D a selection of correlation matrices for this case, and also for the case where modes up to $k\sigma_{\rm p} = 1$ are used, and $\sigma_{\rm p}$ is allowed to float as a further parameter. This makes rather little difference to the precision with which F can be determined (assuming the correctness of the model, of course). The long-

wavelength results are illustrated in Figs 3.5 and 3.6, which show likelihood contours in the (β, F) plane where, at each (β, F) point, the amplitude distribution has been integrated over, but the value of $\sigma_{\rm p}$ is fixed. These are thus not just slices through the likelihood contours, but show the practical joint confidence region on F and β.

3.4.2 Evolutionary effects

In practice, measurable deviations from $F = 1$ will involve working over a significant range of redshifts where one may need to consider evolution of F, β etc. A practical means of dealing with this difficulty will be to divide the survey up into cubes, and use the zero-redshift value of Ω_Λ as a parameter. In each redshift shell, one should form the marginal distribution for Ω_Λ, integrating over other parameters as necessary, before combining all shells to estimate Ω_Λ – in a way that is by construction independent of evolution (an alternative would be to fit a functional form $\beta[z]$). It is the independence of the estimate of Λ on evolution which makes this method so attractive, as most other geometric tests of Λ require assumptions about evolution. In order to apply the Kaiser approximation, we require subsamples that subtend small angles on the sky; cubic subsamples will therefore also have a small $\Delta z/(1+z)$ and so it will be satisfactory to ignore evolution within each cube. The number of modes out to wavenumber $k_{\rm max}$ scales as the box volume, so the number of modes studied is the same whether one large volume is studied, or many small sub-volumes. All that is lost is the signal from the very large-scale modes, but these are a negligible fraction of the total.

3.5 Applications

3.5.1 Galaxies

The next few years should see a number of substantial new galaxy redshift surveys with which this test could be attempted. The *IRAS* PSCz survey of 15000 galaxies (Saunders et al. 1995) will cover essentially 4π sr and will define the density field out to approximately $z \simeq 0.1$. This will not be likely to be so useful for our present purpose

owing both to the high level of shot noise over such volumes and to the low depth, so that $F \lesssim 1.05$ is expected. More promising is the Sloan survey of $\sim 10^6$ galaxies over π sr, which should provide a map free of shot noise out to $z \simeq 0.2$, or an effective volume of about 500^3 $(h^{-1}\mathrm{Mpc})^3$ (Gunn & Weinberg 1995). The AAT 2df facility (Taylor 1995) should provide a survey of slightly greater depth and approximately 1/6 the area, giving errors approximately twice as large as for the Sloan survey.

We have calculated likelihood contours for Ω_Λ for a model survey similar to the 2df survey assuming the same parameter values as in Section 3.4.2, which is shown in Fig. 3.7. The survey is divided into cubes each subtending 13° – side 100 h^{-1} Mpc at $z = 0.2$ – the number of galaxies in each cube being derived from the APM selection function of Baugh & Efstathiou (1993) down to a magnitude limit of $B_\mathrm{J} \leq 19.5$ (a surface density of approximately 150 per square degree). We have assumed for simplicity that β and σ_p' do not vary with redshift; as discussed in Section 3.4.2, this does not affect the precision in Ω_Λ that can be obtained.

We see that the 2df survey is potentially capable of ruling out the high values of $\Omega_\Lambda \simeq 0.8$ that are interesting for inflationary models. Much will depend on the true value of β (which will also be obtained, with exquisite precision). If $\beta \simeq 1$, the rms error on Ω_Λ is about 0.35, but this approximately doubles if $\beta \simeq 0.3$. In this latter case, even the full Sloan survey would barely suffice to detect the flattening effect. This is an ironic outcome, since confusion with the β effect is one of the principal difficulties in measuring the geometrical flattening. Nevertheless, the best hope for a positive geometrical detection comes precisely when the redshift-space distortions are highest.

3.5.2 Quasars

Again, both the Sloan and 2df projects are expected to provide large quasar samples: about 10^5 and 30000 respectively. The larger redshifts probed mean that the anisotropy signal is expected to be significantly larger, and the volumes probed are larger, giving many more modes. However, the quasar samples will be strongly shot-noise limited, reflecting the low space densities of these tracers. In Fig. 3.8 we show results for

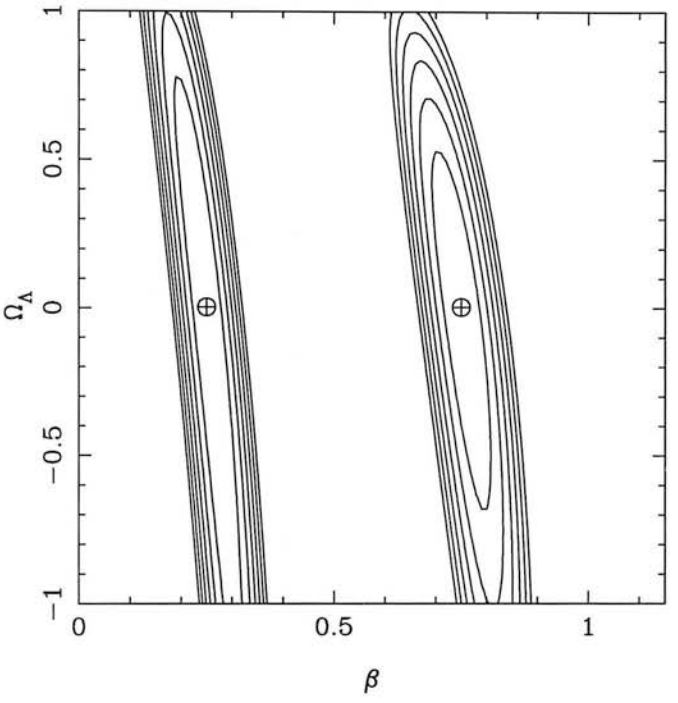

Figure 3.7: Expected contours of likelihood in the $\Omega_\Lambda - \beta$ plane for a 2df-type galaxy redshift survey of 250 000 galaxies to $B_J < 19.5$, obtained by combining cubic subsamples over a range of redshifts. Spatial flatness ($\Omega_m + \Omega_\Lambda = 1$) is assumed. The models shown assume true values for β of 0.25 and 0.75, and $\Omega_\Lambda = 0$, as indicated by the crosses.

30000 quasars over an area of 750 deg^2, with a constant dN/dz over the redshift range $0.3 \leq z \leq 2.2$. For a given value of β, quasars put tighter constraints on a positive Ω_Λ than galaxies. The larger volume, and in particular the higher redshift, more than offsets the increased shot noise, especially given that quasars are observed to be highly clustered at high redshift: the clustering scale-length is claimed to be $r_0 \simeq 6\ h^{-1}$Mpc for quasars with $z < 2.2$ (Shanks & Boyle 1994).

For a particular β, the proposed quasar surveys thus have a strong advantage over the galaxy survey. However, β for quasars may well be smaller. The observed quasar clustering corresponds to $\sigma_8 \simeq 1.2$ at a mean redshift of $z \simeq 1.5$, whereas it is believed that the present value of σ_8 for mass is about $0.6\Omega_m^{-0.5}$; at redshift 1.5 this would be

smaller by 2.5 times the Ω-dependent growth suppression factor (the exact form of which depends on both $\Omega_{\rm m}$ and Ω_Λ). This increases the $\Omega_{\rm m}$ dependence giving roughly

$$\sigma_8(z=1.5) \simeq 0.24\,\Omega_{\rm m}^{-0.8}, \tag{3.43}$$

approximately independent of Λ. If the ratio of the two σ_8 values is taken to measure the bias factor, then quasars are strongly biased at high z and the corresponding β is very low unless $\Omega_{\rm m}$ is extremely low:

$$\beta_{\rm Q}(z=1.5) \simeq 0.2\,\Omega_{\rm m}^{-0.2}. \tag{3.44}$$

However, inspection of Fig. 3.8 shows that, even if β for the quasars is low, they appear to be a better prospect for ruling out high-Ω_Λ models than the low-redshift galaxies. These calculations show that the advantage of having a larger signal at high redshift outweighs the twin disadvantages of higher shot noise and the (anticipated) lower value of β. However, since the datasets are independent, the strongest result would come from combining the likelihoods for the two surveys.

We have assumed that $\sigma_{\rm p}' = 350$ km s^{-1}, as for the galaxy survey. If $\sigma_{\rm p}'$ is as low as 200 km s^{-1} (and this number should certainly decline with redshift, since it scales roughly linearly with σ_8) then $k_{\rm max} = 0.8\ h\,{\rm Mpc}^{-1}$ can be used ($k_{\rm max}\sigma_{\rm p}' = 1.6$). However, the constraints on Ω_Λ are little different from those of Fig. 3.8 – the power-spectrum amplitude falls off towards small scales, hence the additional modes are swamped by the shot noise.

3.6 Summary and Conclusions

We have investigated in some detail the practical applicability of the original suggestion by Alcock & Paczyński (1979) that the cosmological constant might be measured via geometrical distortion of clustering at high redshift. We find that application of this method will be considerably more difficult than previous studies have suggested, for two reasons.

(1) The expectation that $\Omega_{\rm m} \gtrsim 0.2$ limits the likely degree of anisotropy to a factor < 1.3, at least for the spatially flat models popular from inflation.

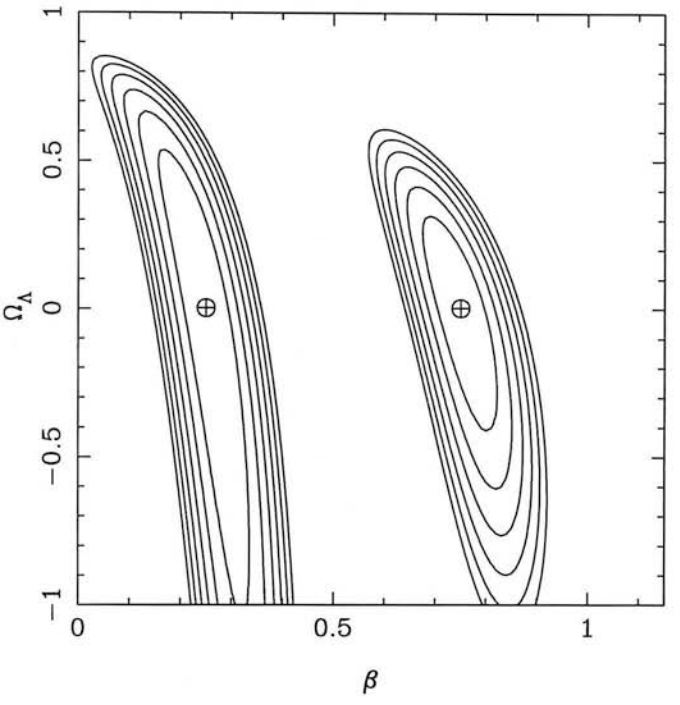

Figure 3.8: Expected contours of likelihood in the $\Omega_\Lambda - \beta$ plane for a 2df-type quasar survey of 30000 quasars with $z < 2.2$. Spatial flatness ($\Omega_m + \Omega_\Lambda = 1$) is assumed. The models shown assume true values for β of 0.25 and 0.75, and $\Omega_\Lambda = 0$, as indicated by the crosses.

(2) Without datasets of very large size and quality, it is difficult to distinguish the geometrical distortion from the redshift-space anisotropies induced by peculiar velocities.

In order to have a chance of distinguishing the two distortions, we have shown that it will be necessary to probe down to scales where the linear analysis of velocity-induced anisotropies is inadequate. We have used an approximate model for such effects to show that datasets likely to become available in the next 5 years stand a good chance of being able to detect Λ. The exact probability of success depends critically on the degree of redshift-space distortion, with stronger signals expected if β is high. For $\beta = 1$, the AAT 2df galaxy redshift survey should be able to detect $\Omega_\Lambda \simeq 0.8$, whereas the full Sloan survey would only just suffice if $\beta \simeq 0.2$. Quasar surveys stand a better chance of

success, provided β_Q is not $\ll \beta_G$. We may expect that surveys being completed now will shortly settle the controversy over at least β_G, giving a clearer idea of how well a geometrical search for Λ will work.

Although we have taken a pessimistic tone, showing that a test for Λ which is purely geometrical will be challenging, there are more optimistic aspects to our conclusions. The forthcoming surveys will give extremely precise measurements of $\beta(z)$, and this will certainly be a useful constraint: a given assumed Λ will then yield $b(z)$, and consistency with the evolution of the clustering amplitude with redshift should allow us to rule out some of these possible bias histories. However, the whole idea of Alcock & Paczyński's original suggestion was to avoid dealing with messy astrophysics, and we have shown that such a 'pure' test is still possible. This test will require a careful understanding of systematics in the redshift surveys, but we believe that it can be made into a reliable method for detecting large vacuum densities. At present, there is no competing method for estimating Λ in a way free from evolutionary uncertainties, so this route will continue to merit detailed scrutiny.

Chapter 4

SPHERICAL HARMONIC ANALYSIS – THEORY

A number of recent redshift surveys have been drawn from objects selected from the *IRAS* Point Source Catalogue (PSC). This catalogue is very even down to a flux limit of $\sim$ 0.6Jy – across a large fraction of the sky the angular selection function can be taken to be constant. Conventional cartesian structure analysis is not well suited to these surveys with their spherical geometry. Clearly, the natural coordinate system in which to analyse such near-all-sky redshift surveys is the spherical polar system, rather than the cartesian system of conventional Fourier analysis. In particular, large scale Fourier modes subtend large angles in these surveys, so that the 'distant observer' approximation cannot be used. The redshift distortion formula of Kaiser (1987) depends on this approximation (Zaroubi & Hoffmann 1996) and hence cannot be used reliably for large scale modes. For example, the 'distant observer' technique of Cole, Fisher & Weinberg (1994, 1995) involved dividing the survey up into sections, losing all the large scale information – *i.e.* throwing away the most reliably linear modes. The natural functions with which to decompose the density field in a near-all-sky survey are spherical harmonics and spherical Bessel functions, which form a complete set of orthonormal basis functions; see Binney & Quinn (1991). In fact, it will be shown that this transform is closely related to a Fourier transform and retains the advantage of a quite clear-cut

division between the linear and nonlinear regimes.

The use of spherical harmonics to analyse galaxy surveys was originally suggested as a 2-dimensional transform for angular surveys (no distance measurements) by Peebles (1973), and was recently resurrected in more recent work. In particular, the method has been used for the analysis of galaxy catalogues drawn from the *IRAS* PSC. Initial work used a 2-dimensional spherical harmonic analysis with the angular PSC catalogue, see Fabbri & Natale (1990) and Scharf *et al.* (1992). The creation of redshift surveys allowed a 3-dimensional analysis; the 2Jy survey was used by Scharf & Lahav (1993), the 1.2Jy by Fisher, Scharf & Lahav (1994) who were the first to measure redshift distortions using a transform involving spherical harmonics. These authors, however, did not use Bessel functions for radial weighting and hence lost the advantages of a power spectrum analysis (see section (1.4.5)).

Some authors have studied alternative radial functions such as logarithmic spherical waves – see Hamilton & Culhane (1996) – although in this case there are problems preserving orthogonality, see Hamilton (1997). Fisher *et al.* (1995) suggest that the Bessel functions could be transformed with a coordinate mapping which causes them to oscillate with a higher frequency where the selection function is highest, analysing the density field more efficiently. However, they point out that this would make the dynamical properties of the Bessel functions more cumbersome.

The full 3-dimensional transform including the Bessel functions was used by Regös & Szalay (1989) to decompose the radial velocity field and by Binney & Quinn (1991) to produce constrained realisations of Gaussian random fields. This transform was used for *IRAS* redshift surveys by Lahav (1993) and Fisher *et al.* (1995) for reconstruction of velocity and density fields. Heavens & Taylor (1995 hereafter HT) used this transform to constrain β and the normalisation amplitude of the real-space power spectrum with the 1.2Jy survey. The work of this chapter and chapter 5 is based on their analysis.

4.1 Formalism and Transform

A scalar field $A(\mathbf{r})$ can be expanded in a spherical transform

$$A(\mathbf{r}) = \sum_{lmn} c_{ln} A_{lmn} j_l(k_{ln} r) Y_{lm}(\Omega), \tag{4.1}$$

with the inverse transform

$$A_{lmn} = c_{ln} \int A(\mathbf{r}) j_l(k_{ln} r) Y^*_{lm}(\Omega)\, \mathrm{d}^3 r. \tag{4.2}$$

The normalisation factor c_{ln} is chosen to appear symmetrically in the transform and its inverse. $j_l(y)$ is the l^{th} order spherical Bessel function, related to ordinary Bessel functions by $j_l(z) = \sqrt{\pi/(2z)} J_{l+\frac{1}{2}}(z)$, and $Y_{lm}(\hat{\mathbf{r}})$ is the Spherical Harmonic function of order l and m, in the angular direction $\hat{\mathbf{r}}$. There are several spherical harmonic conventions, that used by Binney & Tremaine (1987) is adopted here:

$$\begin{aligned} Y_{lm}(\Omega) &= \sqrt{\frac{2l+1}{4\pi} \frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos\theta) \\ &\quad \times \exp(im\phi) \times \begin{cases} (-1)^m & m \geq 0 \\ 1 & m < 0 \end{cases} \end{aligned} \tag{4.3}$$

A boundary condition on a sphere at large $r_{\max}$ specifies the wavenumbers k_{ln}, making the transform discrete rather than continuous. The subscript n is used simply to label the wavenumbers permitted by the boundary conditions for a given l and m, with $n = 1$ denoting the longest wavelength. Note that both c_{ln} and k_{ln} depend on l as well as n – the l subscript in Bessel functions effectively specifies the phase of the radial waves. In this case the boundary condition is chosen to be $j_l'(k_{ln} r_{\max}) = 0$, a Neumann boundary condition (see Arfken 1985 and Jackson 1975). This choice is very convenient because it sets the radial velocity at $r_{\max}$ to zero and hence there is no extra surface term due to redshift distortions. Fisher *et al.* (1995) discuss boundary conditions at some length; they raise objections to this particular choice, but go on to show that it performs very well in simple tests. If the boundary is set at a large enough r so that the selection function is very small, the actual condition chosen shouldn't matter. Using this boundary condition the normalisation becomes

$$c_{ln} = \frac{2k_{ln}^{3/2}}{\sqrt{\pi\{[1/4 + k_{ln}^2 r_{\mathrm{rmax}}^2 - (l+1/2)^2] J_{l+1/2}^2(k_{ln} r_{\mathrm{rmax}})\}}}, \tag{4.4}$$

see appendix F and Abramowitz & Stegun (1972).

For convenience a shorthand will sometimes be used, a single Greek letter subscript representing lmn. For example, the LHS of equation (4.2) could be written as $A_\mu \equiv A_{lmn}$.

This transform is actually the spherical equivalent of a Fourier transform – the kernel functions of both transforms are solutions of the Helmholtz equation,

$$(\nabla^2 + k^2)\psi = 0, \tag{4.5}$$

the different functions arise simply because the equation has been solved in different coordinate systems. The transforms can be related by the Rayleigh equation, an expansion of a plane wave in spherical Bessel functions

$$e^{ikr\cos\theta} = \sum_l (2l+1) i_l P_l(\cos\theta) j_l(kr). \tag{4.6}$$

It is clear from that this that the spherical transform indeed corresponds to a power-spectrum analysis, but with several advantages due to the choice of suitable coordinates, as will be shown.

4.2 Redshift Surveys

In a redshift survey, the observed density field of the galaxies is different from the true underlying density field in several ways. Because the selection function decreases with radial distance, and the mask cuts out some areas of sky, the mean density becomes a function of position. Also, the radial distances are affected by redshift distortions (see chapter 2).

If the mean density of the underlying galaxy density field is ρ_0, and the selection function (angular and radial parts) is $\phi(\mathbf{r})$, then the observed density field can be expressed as

$$\rho'(\mathbf{s}) = \rho_0 \phi(\mathbf{r})(1 + \delta(\mathbf{s})) \tag{4.7}$$

Note that this assumes that the selection criteria are functions of real space; whether

a galaxy of a given luminosity will fall within the flux limit of survey depends on its real-space distance.

The transform of $\rho(\mathbf{s})$ to spherical harmonic space can be generalised to include a weighting function $w(\mathbf{s})$, a function of redshift-space coordinates as it is used to weight the observed galaxies in redshift space. In this case, we consider only radial weighting so $w(\mathbf{s}) \rightarrow w(s)$ so the new transform becomes

$$\tilde{\rho}_{lmn} = c_{ln} \int w(s)\rho(\mathbf{s})\phi(\mathbf{r})j_l(k_{ln}s)Y^*_{lm}(\Omega)\,\mathrm{d}^3s. \tag{4.8}$$

On large scales, coherent infall preserves continuity in the mapping from real to redshift space, $\rho(\mathbf{s})\mathrm{d}^3s = \rho(\mathbf{r})\mathrm{d}^3r$, so the integral can be performed in real space:

$$\tilde{\rho}_{lmn} = c_{ln} \int w(s)\rho(\mathbf{r})\phi(\mathbf{r})j_l(k_{ln}s)Y^*_{lm}(\Omega)\,\mathrm{d}^3r. \tag{4.9}$$

The distortion now only affects two terms, the weighting $w(s)$ and the Bessel function $j_l(k_{ln}s)$. This expression is exact for a coherent velocity field, but says nothing about the actual nature of that field. Equation (4.9) can be expanded to first order in $s - r$, but the velocity field must be related to the density field – in this case using the linear gravitational instability model of section 1.4.5.

The spherical harmonic analysis is particularly suited to analysis of large scale modes, so the corresponding large scale velocity field should be accurately modelled by linear theory. In order to calculate the redshift distortion effect on the transform coefficients, the linear velocity fields must be expressed in terms of the spherical harmonic coefficients of the density field. Using the fact that the transform kernel is an eigenfunction of the Laplacian operator,

$$\nabla^2[j_l(k_{ln}r)Y_{lm}(\Omega)] = -k^2_{ln}j_l(k_{ln}r)Y_{lm}(\Omega), \tag{4.10}$$

equation (1.66) can be used to give the transform of the gravitational potential

$$\phi_{lmn} = -\frac{3\Omega\delta^{\mathrm{m}}_{lmn}}{2k_{ln}}. \tag{4.11}$$

Using this and equation (1.79), and setting $H_0 = 1$, the velocity can be expressed as

$$\mathbf{v} \;=\; -\frac{2}{3}\Omega^{-0.4}\nabla \sum_{lmn} c_{ln}\phi_{lmn}j_l(k_{ln}r)Y_{lm}(\Omega)$$

$$= \Omega^{0.6} \sum_{lmn} c_{ln} \delta^{\rm m}_{lmn} \frac{1}{k_{ln}^2} \nabla [j_l(k_{ln} r) Y_{lm}(\Omega)]$$

$$= \beta \sum_{lmn} c_{ln} \delta_{lmn} \frac{1}{k_{ln}^2} \nabla [j_l(k_{ln} r) Y_{lm}(\Omega)]. \qquad (4.12)$$

The radial component of this is required for redshift distortions, using $\hat{\mathbf{r}} \cdot \nabla = \partial / \partial r$ gives the required relation,

$$U(\mathbf{r}) = v_r(\mathbf{r}) = \hat{\mathbf{r}} \cdot \mathbf{v} = \beta \sum_{lmn} c_{ln} \delta_{lmn} \frac{j_l'(k_{ln} r)}{k_{ln}} Y_{lm}(\Omega). \qquad (4.13)$$

where $j_l'(z) \equiv dj_l/dz$ and $\beta = \Omega_0^{0.6}/b$ as usual. Note that here δ_{lmn} is the transform of the *galaxy* density field. This assumes the linear biasing relationship - $\delta_{lmn} = b\delta^{\rm m}_{lmn}$ see section (1.4.6). Again note that this analysis is ideal for large-scale modes, which are expected to have a linear bias, see Mann, Peacock & Heavens (1997).

If the redshifts are taken in the Local-Group frame of reference, rather than the Microwave Background frame, an extra term should be included (see Hamilton 1997). There is an extra component to the redshift distortion due to the bulk motion of the observer. The distortion is now derived from the radial component of $\mathbf{v} - \mathbf{v}_{\rm LG}$, where $\mathbf{v}$ is the peculiar velocity, and $\mathbf{v}_{\rm LG}$ is the motion of the Local Group relative to the CMB frame, the 'Dipole', see chapter 1. Thus the distortion velocity becomes

$$U(\mathbf{r}) = \hat{\mathbf{r}} \cdot (\mathbf{v} - \mathbf{v}_{\rm LG}) = \beta \sum_{lmn} c_{ln} \delta_{lmn} \frac{j_l'(k_{ln} r)}{k_{ln}} Y_{lm}(\hat{\mathbf{r}}) - \hat{\mathbf{r}} \cdot \mathbf{v}_{\rm LG}. \qquad (4.14)$$

The value of $\mathbf{v}_{\rm LG}$ is derived from other sources and so is not a random variable, hence the extra term only affects the mean values of the transform. As will be seen in chapter 5, this actually has a negligible effect on the parameters derived using this model. However, the effect on the 'Dipole' technique for measuring β can be more severe, see chapter 1 and Kaiser (1987).

By equation (1.38), the density can be expressed in terms of a mean ρ_0 and a fractional overdensity $\delta(\mathbf{r})$,

$$\rho(\mathbf{r}) = \rho_0 [1 + \delta(\mathbf{r})]. \qquad (4.15)$$

Using this, equation (4.9) is expanded to first order in $(s - r)$,

$$
\begin{aligned}
\tilde{\rho}_{lmn} &= c_{ln} \int \rho_0 \, \phi(\mathbf{r}) w(s) j_l(k_{ln} s) Y^*_{lm}(\Omega) \, \mathrm{d}^3 r \\
&+ c_{ln} \int \rho_0 \, \delta(\mathbf{r}) \phi(\mathbf{r}) w(s) j_l(k_{ln} s) Y^*_{lm}(\Omega) \, \mathrm{d}^3 r \\
&= c_{ln} \int \rho_0 \, \phi(\mathbf{r}) w(r) j_l(k_{ln} r) Y^*_{lm}(\Omega) \, \mathrm{d}^3 r \\
&+ c_{ln} \int \rho_0 \, \phi(\mathbf{r}) (s - r) \frac{\mathrm{d}}{\mathrm{d}r} [w(r) j_l(k_{ln} r)] \, Y^*_{lm}(\Omega) \, \mathrm{d}^3 r \\
&+ c_{ln} \int \rho_0 \, \delta(\mathbf{r}) \phi(\mathbf{r}) w(r) j_l(k_{ln} r) Y^*_{lm}(\Omega) \, \mathrm{d}^3 r \\
&+ c_{ln} \int \rho_0 \, \phi(\mathbf{r}) \delta(\mathbf{r}) (s - r) \frac{\mathrm{d}}{\mathrm{d}r} [w(r) j_l(k_{ln} r)] Y^*_{lm}(\Omega) \, \mathrm{d}^3 r. \qquad (4.16)
\end{aligned}
$$

The final term is of order $\delta(\mathbf{r})(s - r)$, so can be ignored for large scale modes with $\delta(\mathbf{r}) \ll 1$. Using equation (4.14) to express $(s - r)$ in terms of the density field, and expanding $\delta(\mathbf{r})$ with equation (4.1),

$$
\begin{aligned}
\tilde{\rho}_{lmn} &= c_{ln} \int \rho_0 \, \phi(\mathbf{r}) w(r) j_l(k_{ln} r) Y^*_{lm}(\Omega) \, \mathrm{d}^3 r \\
&+ c_{ln} \int \rho_0 \, \phi(\mathbf{r}) \, \beta \sum_{l'm'n'} c_{l'n'} \, \delta_{l'm'n'} \frac{j'_{l'}(k_{l'n'} r)}{k_{l'n'}} Y_{l'm'}(\Omega) \\
&\quad \times \frac{\mathrm{d}}{\mathrm{d}r} [w(r) j_l(k_{ln} r)] \, Y^*_{lm}(\Omega) \, \mathrm{d}^3 r \\
&+ c_{ln} \int \rho_0 \sum_{l'm'n'} c_{l'n'} \, \delta_{l'm'n'} \, j_{l'}(k_{l'n'} r) Y_{l'm'}(\Omega) \\
&\quad \times \phi(\mathbf{r}) w(r) j_l(k_{ln} r) Y^*_{lm}(\Omega) \, \mathrm{d}^3 r \\
&- c_{ln} \int \rho_0 \, \phi(\mathbf{r}) \, \hat{\mathbf{r}} \cdot \mathbf{v}_{\mathrm{LG}} \frac{\mathrm{d}}{\mathrm{d}r} [w(r) j_l(k_{ln} r)] \, Y^*_{lm}(\Omega) \, \mathrm{d}^3 r \qquad (4.17)
\end{aligned}
$$

The varying selection, both the radial part $\phi(r)$ and the mask $\mathrm{M}(\Omega)$, lead to additional apparent structure, as do redshift distortions. From equation (4.17), the effect is to convolve the transform of the underlying density field with a window function.

$$
\tilde{\rho}_{lmn} = (\rho_0)_{lmn} + \sum_{l'm'n'} (\Phi^{mm'}_{ll'nn'} + \beta V^{mm'}_{ll'nn'}) \delta_{l'm'n'}, \qquad (4.18)
$$

where the mean of the coefficients is no longer zero in general,

$$
\begin{aligned}
(\rho_0)_{lmn} &= c_{ln} \int \rho_0 \phi(r) w(r) j_l(k_{ln} r) Y_{lm}(\Omega) \mathrm{d}^3 r \\
&- c_{ln} \int \rho_0 \, \phi(\mathbf{r}) \, \hat{\mathbf{r}} \cdot \mathbf{v}_{\mathrm{LG}} \frac{\mathrm{d}}{\mathrm{d}r} [w(r) j_l(k_{ln} r)] \, Y^*_{lm}(\Omega) \, \mathrm{d}^3 r, \qquad (4.19)
\end{aligned}
$$

where the second term is the Local-Group correction. The mixing matrices due to the selection function and (for the V matrices) redshift distortion are

$$\Phi^{mm'}_{ll'nn'} \equiv c_{ln}c_{l'n'}\int \rho_0\phi(\mathbf{r})w(r) \times j_l(k_{ln}r)Y_{lm}(\Omega)j_{l'}(k_{l'n'}r)Y^*_{l'm'}(\Omega)\mathrm{d}^3r, \quad (4.20)$$

and

$$V^{mm'}_{ll'nn'} \equiv \frac{c_{ln}c_{l'n'}}{k^2_{ln}}\int \rho_0\phi(\mathbf{r})\frac{\mathrm{d}}{\mathrm{d}r}\left[w(r)j_l(k_{ln}r)\right] \times Y_{lm}(\Omega)j'_{l'}(k_{l'n'}r)Y^*_{l'm'}(\Omega)\mathrm{d}^3r. \quad (4.21)$$

These can be used to write the transform of the observed fractional overdensity field, $D_\mu \equiv (\tilde{\rho}_{lmn} - (\tilde{\rho}_0)_{lmn})/\rho_0$, in terms of the underlying one

$$D_\mu = \sum_\nu \Lambda^\nu_\mu \delta_\nu, \quad (4.22)$$

with the combined mixing matrix

$$\Lambda^\nu_\mu \equiv \Phi^\nu_\mu + \beta V^\nu_\mu. \quad (4.23)$$

Equation (4.22) deals correctly with a general (not just radial) selection function, and with the effect of linear theory redshift distortions. Note that the V matrices show that the redshift distortions lead to an *additional* mixing of modes, in addition to the basic selection function effect described by the Φ matrix. This is clearly different from the Kaiser equation, which predicts a simple angle-dependent boost to the spectrum. It does, however, have much in common with equation (2.19), the cartesian formula of Zaroubi & Hoffmann (1996). Both equations describe mode-mixing effects due to selection functions and redshift distortions. The analogy will become clearer when the covariance matrix of the transform coefficients is calculated.

4.2.1 Small–Scale Velocity Structure

Although the spherical harmonic analysis is particularly useful for analysing large-scale linear modes, there is still some risk of contamination by the nonlinear velocity field from smaller scales, particularly if the mixing effect is strong. As mentioned in Chapter 2, the first order effect of nonlinear velocities can be modelled as an incoherent component

in the velocity field, destroying the one-to-one mapping of linear theory. This is often regarded as the effect of random motions in virialised cluster, the 'Fingers of God', which cause clusters to stretch out radially in redshift space. The randomisation smears out structure along the line of sight, reducing radial power, *i.e.* acting in the opposite sense to linear infall. If linear models are used, the apparent value of β can be reduced by this effect. As noted earlier, this simple randomisation model can also correct for other nonlinear effects, particularly if the nonlinearity is mild.

As suggested by Peacock (1992), a simple Gaussian was used to model the random motions. Note that this assigns the same dispersion to every galaxy, regardless of position and local density – see the detailed discussion in chapter 2. It has, however, proved to be accurate for the mildly nonlinear regime, and the spherical harmonic-Bessel method is designed to avoid nonlinearity if at all possible. Any contamination should be very small and accurately modelled by this simple technique.

The redshift space coordinate $\mathbf{s}(\mathbf{r})$ is scattered by the addition of a random variable ε:

$$\mathbf{s}(\mathbf{r}) \longrightarrow \mathbf{s}'(\mathbf{r}) = \mathbf{s}(\mathbf{r}) + \varepsilon(\mathbf{r}). \tag{4.24}$$

The mapping is no longer on-to-one, so we can only find a mean observed mode for a given density field. The new observed mode is taken to be the ensemble average of the observed modes of equation (4.22) with the new, additional distortion:

$$D'(\mathbf{s}) = \langle D[\mathbf{s}'(\mathbf{s})] \rangle_\varepsilon. \tag{4.25}$$

This does not interfere with subsequent ensemble averaging, as the random scatter is assumed to be independent of the particular realisation of the density field. Specifically, the covariance matrix in section (4.3) is given by

$$C_{\alpha\beta} = \langle D_\alpha D_\beta \rangle. \tag{4.26}$$

Ensemble averaging over realisations of ε, $\langle D_\alpha D_\beta \rangle \rightarrow \langle \langle D_\alpha D_\beta \rangle \rangle_\varepsilon$. Because of the independence of the scatter vector, this can be written as

$$\langle \langle D_\alpha D_\beta \rangle \rangle_\varepsilon = \langle \langle D_\alpha \rangle_\varepsilon \langle D_\beta \rangle_{\varepsilon'} \rangle. \tag{4.27}$$

Hence only the formula for $D'(\mathbf{s})$ is needed.

Transforming using the average over realisations gives

$$\begin{aligned} D'_{lmn} &= c_{ln}\langle \int \phi(\mathbf{r})\delta(\mathbf{r})w(\mathbf{s}')j_l(k_{ln}s')Y_{lm}(\Omega_{s'})\mathrm{d}^3r\rangle_\varepsilon, \\ &= c_{ln}\int \phi(\mathbf{r})\delta(\mathbf{r})\int p(\mathbf{s}-\mathbf{s}')w(\mathbf{s}')j_l(k_{ln}s')Y_{lm}(\Omega_{s'})\mathrm{d}^3s'\mathrm{d}^3r, \\ &= c_{ln}\int \phi(\mathbf{r})w(\mathbf{s})\delta(\mathbf{r})\int p(\mathbf{s}-\mathbf{s}')\frac{w(\mathbf{s}')}{w(\mathbf{s})}j_l(k_{ln}s')Y_{lm}(\Omega_{s'})\mathrm{d}^3s'\mathrm{d}^3r, \\ &= c_{ln}\int \phi(\mathbf{r})w(\mathbf{s})\delta(\mathbf{r}) \\ &\times \iint p(\mathbf{x}-\mathbf{s}')\frac{w(\mathbf{s}')}{w(\mathbf{x})}j_l(k_{ln}s')Y_{lm}(\Omega_{s'})\delta_{\mathrm{D}}(\mathbf{x}-\mathbf{s})\mathrm{d}^3x\mathrm{d}^3s'\mathrm{d}^3r. \end{aligned} \tag{4.28}$$

Using the orthogonality and completeness relations, the Dirac delta function can be expanded in terms of spherical wave functions,

$$\delta_{\mathrm{D}}(\mathbf{x}-\mathbf{s}) = c^2_{l'n'}\sum_{l'm'n'} j_l(k_{l'n'}x)Y_{l'm'}(\Omega_x)j_l(k_{l'n'}s)Y^*_{l'm'}(\Omega_s) \tag{4.29}$$

Substituting this into equation (4.28), the formula can be split into a convolution of the original density transform with a scattering matrix

$$D'_{lmn} = \sum_{l'n'm'} S^{mm'}_{ll'nn'}D_{l'm'n'}; \tag{4.30}$$

the scattering matrix is given by

$$\begin{aligned} S^{mm'}_{ll'nn'} \equiv\ & c_{ln}c_{l'n'}\iint p(\mathbf{x}-\mathbf{s}')j_l(k_{ln}s')Y^*_{lm}(\Omega_{s'}) \\ & \times j_{l'}(k_{l'n'}x)Y_{l'm'}(\Omega_x)s'^2\mathrm{d}s'\ \mathrm{d}\Omega'_s\, x^2\mathrm{d}x\,\mathrm{d}\Omega_x\,. \end{aligned} \tag{4.31}$$

The radial component of the random variable $\varepsilon(\mathbf{r})$ is assigned a Gaussian distribution with mean $\langle\varepsilon(\mathbf{r})\rangle = 0$ and variance $\langle\varepsilon_i(\mathbf{r}_1)\varepsilon_j(\mathbf{r}_2)\rangle = \sigma^2\delta_{ij}.\delta_D(\mathbf{r}_1-\mathbf{r}_2)$. Redshift distortion effects are purely radial in nature, so there is no angular component. Hence the 3-dimensional scattering probability can be decomposed into a radial Gaussian and an angular delta function,

$$p(\mathbf{r}-\mathbf{y}) = \left[\frac{e^{-(r-y)^2/(2\sigma^2)}}{\sqrt{2\pi}\sigma}\right]\delta_D(\Omega_r-\Omega_y). \tag{4.32}$$

With this distribution, the scattering matrix simplifies to

$$\begin{aligned} S^{mm'}_{ll'nn'} &= c_{ln}c_{l'n'}\delta^{\mathrm{K}}_{ll'}\delta^{\mathrm{K}}_{mm'}\int\int \\ &\times\frac{e^{-(r-y)^2/(2\sigma^2)}}{\sqrt{2\pi}\sigma}j_l(k_{ln}r)j_{l'}(k_{l'n'}y)r\mathrm{d}r\,y\mathrm{d}y, \end{aligned} \tag{4.33}$$

which can be included in the transformations to derive the observed density field,

$$D'_{lmn} = S^{mm''}_{ll''nn''}(\Phi^{m''m'}_{l''l'n''n'} + \beta V^{m''m'}_{l''l'n''n'})\delta_{l'm'n'}, \tag{4.34}$$

where summation is implied.

This simple analysis can also model random errors in redshift measurements – the additional distance error is simply added in quadrature.

4.2.2 Division into Radial and Angular Selection

When dealing with selection functions, the spherical harmonic techniques have a major advantage over methods using other coordinate systems. For most redshift surveys, the selection function separates quite naturally into a radial part and an angular part.

$$\phi(\mathbf{r}) = M(\Omega)\phi(r), \tag{4.35}$$

Where $M(\Omega)$ is the mask function and $\phi(r)$ is the radial selection function. This convenient separability arises because the coordinates are so appropriate for the selection criteria used in most surveys. The radial part is often a falling function of r in a flux limited survey, or a sharp cutoff in r for a volume limit. The angular part is usually independent of r, in particular the survey is often taken to a fixed flux or radius limit over a particular patch of the sky, creating an angular selection function

$$M(\Omega) = \begin{cases} 1 & \text{if observable area of sky,} \\ 0 & \text{otherwise} \end{cases} \tag{4.36}$$

For near-all-sky surveys the masked region $M(\Omega) = 0$ is quite small, minimising the mixing effects of the angular part of the selection function. In particular the 'zone of avoidance', the region obscured by the galactic plane, is masked, leading to an large quadrupole in the observed modes as well as a mixing effect.

The mixing matrices of equation (4.20) and equation (4.21) can be factored into angular and radial parts, becoming

$$\begin{aligned} \Phi^{mm'}_{ll'nn'} &= W^{mm'}_{ll'}\Phi^{nn'}_{ll'} \\ V^{mm'}_{ll'nn'} &= W^{mm'}_{ll'}V^{nn'}_{ll'} \end{aligned} \tag{4.37}$$

with no sum on l and l', and the separate matrices are

$$\Phi_{ll'}^{nn'} \equiv c_{ln}c_{l'n'}\int \rho_0(r)w(r)j_{l'}(k_{l'n'}r)j_l(k_{ln}r)r^2\mathrm{d}r \tag{4.38}$$

$$V_{nn'}^{ll'} \equiv \frac{c_{ln}c_{l'n'}}{k_{l'n'}^2}\int \rho_0(r)\frac{d}{dr}\left[w(r)j_{l'}(k_{l'n'}r)\right]j_l'(k_{ln}r)r^2\mathrm{d}r \tag{4.39}$$

$$W_{ll'}^{mm'} \equiv \int_\Omega Y_{l'm'}\,M(\Omega)\,Y_{lm}^*\,\mathrm{d}\Omega. \tag{4.40}$$

The complete transform can now be expressed as

$$D'_{lmn} = S_{ll''nn''}^{mm''}W_{ll'}^{mm'}(\Phi_{l''l'}^{n''n'} + \beta V_{l''l'}^{n''n'})\delta_{l'm'n'}, \tag{4.41}$$

where summation is implied.

For a truly all-sky survey, with no mask – $M(\Omega) = 1$ everywhere – the orthogonality relations on the spherical harmonics diagonalise the angular matrices: $W_{ll'}^{mm'} = \delta_{ll'}^{\mathrm{K}}\delta_{mm'}^{\mathrm{K}}$. Hence there is only mixing between modes of the same l and m, a considerable simplification. Even without this simplification, the covariance matrices are 4D rather than the huge 6D matrices which arise in the general Zaroubi & Hoffmann (1996) formalism for the Fourier case (equation (2.19)), where all modes in any cartesian direction mix with each other.

4.3 Statistics

Theories of the origin of structure do not specify the precise details of the density field, rather they predict its statistical properties. Inflation theory (section 1.2.2)predicts that the primordial density fluctuations are in the form of a Gaussian random field, the statistical properties of which are fully specified by the power spectrum – see section 1.4.2. The Central Limit Theorem also lends weight to the belief that the initial fluctuations were Gaussian.

4.3.1 Covariance matrix terms – real and imaginary parts

The observed modes of the transform are complex numbers, so we can treat the real and imaginary parts as separate data. The (real, symmetric) covariance matrix for this data

is composed of four sections; a real part, an imaginary part, and two crossover parts which are transposes of each other (covariances between real and imaginary modes)

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{\Re e\Re e} & \mathbf{C}_{\Re e\Im m} \\ \mathbf{C}_{\Im m\Re e} & \mathbf{C}_{\Im m\Im m} \end{pmatrix} \tag{4.42}$$

The overdensity field transform from equation (4.41) is (in matrix form)

$$\mathbf{D} = \mathbf{SW}(\Phi + \beta\mathbf{V})\boldsymbol{\delta}. \tag{4.43}$$

In general, $\mathbf{W}$ is the only transition matrix which can have an imaginary part (see equations (4.38) to (4.40)), so splitting all complex matrices

$$\Re e\mathbf{D} + i\Im m\mathbf{D} = (\Re e\mathbf{W} + i\Im m\mathbf{W})(\Phi + \beta\mathbf{V})(\Re e\boldsymbol{\delta} + i\Im m\boldsymbol{\delta}) \tag{4.44}$$

$$\Re e\mathbf{D} = \Re e\mathbf{W}(\Phi + \beta\mathbf{V})\Re e\boldsymbol{\delta} - \Im m\mathbf{W}(\Phi + \beta\mathbf{V})\Im m\boldsymbol{\delta} \tag{4.45}$$

$$\Im m\mathbf{D} = \Im m\mathbf{W}(\Phi + \beta\mathbf{V})\Re e\boldsymbol{\delta} + \Re e\mathbf{W}(\Phi + \beta\mathbf{V})\Im m\boldsymbol{\delta}. \tag{4.46}$$

The imaginary part of the window function introduces mixing between real and imaginary parts of the density transform vector $\boldsymbol{\delta}$. If these parts are zero, *e.g.* when the mask function $M(\Omega)$ is azimuthally symmetric, the off-diagonal quarters of the covariance matrix in equation (4.42) are zero, simplifying (and speeding up) the analysis.

The covariances of the modes of the underlying density field are given by the power spectrum; for modes with $m \neq 0$

$$\langle \Re e\delta_{lmn}\Re e\delta_{l'm'n'}\rangle = \langle \Im m\delta_{lmn}\Im m\delta_{l'm'n'}\rangle = \frac{1}{2}P(k_{ln})\delta^{\mathrm{K}}_{ll'}\delta^{\mathrm{K}}_{mm'}\delta^{\mathrm{K}}_{nn'}. \tag{4.47}$$

Modes with $m = 0$ are entirely real because Y_{l0} spherical harmonics are real as is the density field. This means that all the power in these modes is in the real part

$$\langle \Re e\delta_{l0n}\Re e\delta_{l'm'n'}\rangle = P(k_{ln})\delta^{\mathrm{K}}_{ll'}\delta^{\mathrm{K}}_{0m'}\delta^{\mathrm{K}}_{nn'}, \tag{4.48}$$

and

$$\langle \Im m\delta_{l0n}\Im m\delta_{l'm'n'}\rangle = 0. \tag{4.49}$$

The cross terms are independent for a Gaussian field

$$\langle \Re e\delta_{lmn}\Im m\delta_{l'm'n'}\rangle = \langle \Im m\delta_{lmn}\Re e\delta_{l'm'n'}\rangle = 0. \tag{4.50}$$

The covariance matrix elements for the *observed* field are thus given by

$$\begin{aligned}\langle \Re eD_\mu \Re eD_\nu \rangle &= \frac{1}{2}\sum_\alpha [\Re eW_\mu(\Phi_\mu^\alpha + \beta V_\mu^\alpha)\Re eW_\nu(\Phi_\nu^\alpha + \beta V_\nu^\alpha) \\ &+ \Im mW_\mu(\Phi_\mu^\alpha + \beta V_\mu^\alpha)\Im mW_\nu(\Phi_\nu^\alpha + \beta V_\nu^\alpha)]P(k_\alpha) \\ &= \langle \Im mD_\mu \Im mD_\nu \rangle, \end{aligned} \tag{4.51}$$

with cross terms from the off-diagonal blocks

$$\begin{aligned}\langle \Re eD_\mu \Im mD_\nu \rangle &= \frac{1}{2}\sum_\alpha [\Re eW_\mu(\Phi_\mu^\alpha + \beta V_\mu^\alpha)\Im mW_\nu(\Phi_\nu^\alpha + \beta V_\nu^\alpha) \\ &- \Im mW_\mu(\Phi_\mu^\alpha + \beta V_\mu^\alpha)\Re eW_\nu(\Phi_\nu^\alpha + \beta V_\nu^\alpha)]P(k_\alpha) \\ &= -\langle \Im mD_\mu \Re eD_\nu \rangle, \end{aligned} \tag{4.52}$$

4.3.2 Shot noise

The discreteness of galaxies adds extra apparent structure to the galaxy density field. It is assumed that the galaxies randomly Poisson sample a continuous field, and the effect on the observed modes can be calculated using the technique of Peebles (1980, §41). The density field of discrete point particles can be expressed as a set of delta functions,

$$\rho(\mathbf{s}) = \sum_i \delta_{\mathrm{D}}(\mathbf{s}(\mathbf{r},\mathbf{v}) - \mathbf{r}_i)\delta_{\mathrm{D}}(\mathbf{v} - \mathbf{v}_i), \tag{4.53}$$

translating the transform into a summation over the particles:

$$\tilde{\rho}_{lmn} = c_{ln}\sum_i w(s_i)\phi(\mathbf{s}_i)j_l(k_{ln}s_i)Y_{lm}^*(\theta_i,\phi_i). \tag{4.54}$$

The survey is split into infinitesimal cells, each containing either 0 or 1 galaxies – the probability of a cell containing more than one is vanishingly small. The transform becomes a sum over cells,

$$\tilde{\rho}_{lmn} = c_{ln}\sum_{\text{cell}\,i} n_i w(s_i)\phi(\mathbf{s}_i)j_l(k_{ln}s_i)Y_{lm}^*(\theta_i,\phi_i), \tag{4.55}$$

where $n_i = 1$ if the cell contains a galaxy, 0 otherwise.

As discussed in the section (4.3.1), the real and imaginary parts of the transform coefficients can be treated as separate data, creating a real, symmetric covariance matrix.

The elements of the matrix are given by

$$\langle D_{\mu a} D_{\nu b} \rangle = \langle \rho_{\mu a} \rho_{\nu b} \rangle - \langle \rho_{\mu a} \rangle \langle \rho_{\nu b} \rangle, \tag{4.56}$$

with $a, b = 0, 1$ signifying real or imaginary parts of modes or spherical harmonics. Using equation (4.55), the first term on the right-hand side becomes

$$\begin{aligned} \langle \rho_{\mu a} \rho_{\nu b} \rangle &= c_{ln} c_{l'n'} \sum_{i,j} \langle n_i n_j \rangle w_i j_l(k_{ln} r_i) Y_{lma}(\Omega_i) w_j j_{l'}(k_{l'n'} r_j) Y_{l'm'b}(\Omega_j) \\ &= c_{ln} c_{l'n'} \sum_{i} \langle n_i \rangle w_i^2 j_l(k_{ln} r_i) Y_{lma}(\Omega_i) j_{l'}(k_{l'n'} r_i) Y_{l'm'b}(\Omega_i) \\ &+ c_{ln} c_{l'n'} \sum_{\substack{i,j \\ i \neq j}} \langle n_i n_j \rangle w_i j_l(k_{ln} r_i) Y_{lma}(\Omega_i) w_j j_{l'}(k_{l'n'} r_j) Y_{l'm'b}(\Omega_j) \end{aligned} \tag{4.57}$$

From the definition of the correlation function,

$$\langle n_i n_j \rangle = \langle n_i \rangle \langle n_j \rangle [1 + \xi(\mathbf{r}_i - \mathbf{r}_j)] \tag{4.58}$$

so the second term in the rhs of equation (4.57) splits into two, with a signal (correlation) part and a mean term which cancels with the second term in equation (4.56). Hence the full covariance matrix, including noise, can be expressed as

$$C_{\mu\nu} = \langle D_\mu D_\nu \rangle = \frac{1}{2} \sum_\alpha (\Phi_\mu^\alpha + \beta V_\mu^\alpha)(\Phi_\nu^\alpha + \beta V_\nu^\alpha) P(k_\alpha) + \Lambda_{\mu\nu}^0, \tag{4.59}$$

with noise matrix

$$\begin{aligned} \Lambda_{\mu\nu}^0 &= c_{ln} c_{l'n'} \int \rho_0(r) w^2(r) j_l(k_{ln} r) j_{l'}(k_{l'n'} r) r^2 \mathrm{d}r \times \\ &\quad \int \mathcal{P}_\mu(Y_{lm}(\Omega)) M(\Omega) \mathcal{P}_\nu(Y_{l'm'}^*(\Omega)) \mathrm{d}\Omega. \end{aligned} \tag{4.60}$$

The tesseral harmonics $\mathcal{R}e$ or $\mathcal{I}m[Y_{lm}(\Omega)]$ may be written, for example, $\mathcal{R}e[Y_{lm}(\Omega)] = \frac{1}{2}[Y_{lm}(\Omega) + Y_{lm}^*(\Omega)]$ – see Appendix (E.2).

4.3.3 Likelihood analysis

As mentioned in section (1.4.2), it seems reasonable to assume that the large-scale *linear* density field is a Gaussian random field, the statistical properties of which are entirely characterised by the power spectrum. Following section (1.4.2) the probability

distribution of the underlying density field can be expressed in terms of its harmonic expansion

$$P[\delta] = (\prod_{lmn} P(k_{ln}))^{-1/2} \exp\left(-\frac{1}{2}\sum_{lmn}\frac{\delta_{lmn}\delta^*_{lmn}}{P(k_{ln})}\right), \tag{4.61}$$

where the real and imaginary parts are independent Gaussian variables with zero mean. Because they are statistically independent, the modes are uncorrelated and the covariance matrix is diagonal, just the power spectrum $P(k_{ln})$ (with determinant $\prod_{lmn} P(k_{ln})$), and hence the summation is over only one set of (lmn). The independence of the real and imaginary parts can be used to split the distribution into two

$$\begin{aligned}
P[\delta] &= P[\mathcal{R}e(\delta_{lmn})]\ P[\mathcal{I}m(\delta_{lmn})], \\
P[\mathcal{R}e(\delta_{lmn})] &= (\det A_{\mathcal{R}e})^{-1/2} \exp\left(-\sum_{lmn}\frac{\mathcal{R}e(\delta_{lmn})^2}{2P(k_{ln})}\right), \\
P[\mathcal{I}m(\delta_{lmn})] &= (\det A_{\mathcal{I}m})^{-1/2} \exp\left(-\sum_{lmn}\frac{\mathcal{I}m(\delta_{lmn})^2}{2P(k_{ln})}\right),
\end{aligned} \tag{4.62}$$

except for the entirely real $m = 0$ modes:

$$\begin{aligned}
P[\delta] &= P[\mathcal{R}e(\delta_{lmn})]\ P[\mathcal{I}m(\delta_{lmn})], \\
P[\mathcal{R}e(\delta_{lmn})] &= (\det A_{\mathcal{R}e})^{-1/2} \exp\left(-\sum_{lmn}\frac{\mathcal{R}e(\delta_{lmn})^2}{2P(k_{ln})}\right), \\
P[\mathcal{I}m(\delta_{lmn})] &= \delta_{\mathrm{D}}(\mathcal{I}m(\delta_{lmn}) - 0).
\end{aligned} \tag{4.63}$$

We do not directly observe the underlying density field, but rather the distorted modes modelled by equation (4.59). Fortunately, a field always retains its Gaussianity under a linear transformation[1], so the Gaussian assumption is no less valid. If correlations are introduced between real and imaginary modes, the probability can no longer be split into separate parts,

$$\begin{aligned}
\mathcal{L}[D|\beta, P(k_{ln})] &= (\det \mathbf{C})^{-1/2} \\
&\times \exp\left(-\frac{1}{2}\sum_{\mu a \nu b}(D_{\mu a})(C^{-1})_{\mu a \nu b}(D_{\nu b})\right)
\end{aligned} \tag{4.64}$$

where, as before, subscripts $a, b = 0, 1$ are used to identify real and imaginary parts.

[1] sums of Gaussian variables are still Gaussian variables.

The model outlined in equation (4.59) includes the parameter β and the function $P(k)$, the power spectrum. The power spectrum can be parameterised in a number of ways; in a very simple case the shape can be fixed and only the amplitude fitted as a free parameter, giving a simple two parameter model. Alternatively, a more complicated parameterisation can be used to fit the shape of the power spectrum and β simultaneously. Both of these possibilities are explored in this thesis.

The degeneracy between β and the power spectrum amplitude is broken by the phase information stored in the harmonic modes. A change in amplitude changes the expected variance of all modes equally, whereas redshift distortions increase radial power only. The actual details of how the distortion is measured is not as transparent with this technique as it is with the methods which explicitly compare clustering parallel with and perpendicular to the line of sight – *e.g.* Cole *et al.* (1994). In detail, the anisotropy is measured by differences in measured $P(k_{ln})$ as a function of l at a fixed k, higher l values are less affected by distortion. Clearly change in amplitude would affect all modes of fixed k equally. The l dependence can be understood as follows: if l is high, the Bessel function has a flat $j_l(k_{ln}r) \simeq 0$ portion at small r before it starts to oscillate – the higher l, the longer this 'lag'. Thus the contribution to the distortion integral in equation (4.40) is smaller for higher l. Essentially, high-l modes contain a lot of angular structure (high order of $Y_{lm}(\Omega)$) at high r, where the effect of distortion is weaker – only the second factor of the Jacobian in equation (2.7) is significant.

Perhaps the simplest way to understand how the distortions can be disentangled from changes in amplitude is to realise that the spherical harmonic transform of equation (4.2) is invertible, *i.e.* it contains all the information about the observed density field, including phase information. Hence a model with the wrong degree of anisotropy and inhomogeneity should not fit the data so well.

4.3.4 Optimal Weighting of the data

In equation (4.8), the transform of the density field, there is an weighting function, $w(\mathbf{r})$, which can be chosen arbitrarily. It is possible to find a function which optimally

weights the galaxies for the extraction of a statistic of interest, *e.g.* the parameter β or the function $P(k)$. This weighting scheme will give the smallest error bar on a parameter in the likelihood analysis. In this case we consider only radial weighting: $w(\mathbf{r}) = w(r)$.

It should be noted that although approximations are made to derive these weighting schemes, there is complete freedom to weight the galaxies in any way. The schemes can be assessed by their effect on measurements of statistics.

The variance of a statistic must be minimised with respect to a weighting function, so the functional derivative must be zero,

$$\frac{\delta}{\delta w(y)} \mathrm{Var}[f(r)] = 0, \tag{4.65}$$

where f is the statistic being measured. The variance of f can be approximated from

$$\mathrm{Var}[f(r)] \simeq -\left\langle \frac{\delta^2 \ln \mathcal{L}[f(y)]}{\delta f(r)^2} \right\rangle^{-1}, \tag{4.66}$$

where $-\ln \mathcal{L}$ is called the likelihood functional; this is just the likelihood function if f is a discrete parameter. The quantity $-1/\mathrm{Var}[f]$ is known as the amount of information. If f is a discrete parameter $-\ln \mathcal{L}$ is simply the likelihood function equation and the rhs of equation 4.66 is the diagonal term of the Fisher information matrix – see chapter 6 and Tegmark, Taylor & Heavens (1997). This formula for the variance is exact if the likelihood of the parameter is Gaussian and there are no correlations with other parameters. Applying this expression to the Gaussian random field model, we find (see equation (4.64))

$$\mathrm{Var}[f(r)] \simeq 2 \sum_{\mu} \{[(\mathbf{C}')^{-1}\mathbf{C}]^2\}_{\mu\mu}, \tag{4.67}$$

where $\mathbf{C}$ is the covariance matrix and $\mathbf{C}' \equiv \delta\mathbf{C}/\delta f$ is the functional derivative of $\mathbf{C}$ with respect to $f(r)$. The effect of correlations between observed modes – *i.e.* off-diagonal terms – has been ignored.

Ignoring redshift distortions, an optimal weighting scheme for measuring the power spectrum can be derived using the above formalism. After this the analysis will be generalised to include distortion effects.

Setting $\beta = 0$ in equation (4.59), and assuming $P(k)$ is constant over the range of interest, a diagonal term of the covariance matrix is

$$C_{ln} \approx \int w^2(r)\phi^2(r)\left(\frac{1}{\phi(r)} + P(k_{ln})\right) j_l^2(k_{ln}r)\,\mathrm{d}^3r, \tag{4.68}$$

where the orthogonality relations (appendix F) are used, reducing the summation over modes.

Using equation (4.67), the fractional variance of the power spectrum is

$$\frac{\mathrm{Var}P(k_{ln})}{P^2(k_{ln})} \approx \frac{2\int w^4(r)\phi^4(r)\left|1+\frac{1}{\phi(r)P(k_{ln})}\right|^2 j_l^4(k_{ln}r)\,\mathrm{d}^3r}{|\int w^2(r)\phi^2(r)j_l^2(k_{ln}r)\,\mathrm{d}^3r|^2}. \tag{4.69}$$

Taking the functional derivative with respect to the weighting function, $w(y)$, and using the functional relation $\delta w(\mathbf{x})/\delta w(\mathbf{y}) = \delta_D(\mathbf{x}-\mathbf{y})$, equation 4.65 becomes

$$\frac{\delta}{\delta w}\ln\frac{\mathrm{Var}P}{P^2} \approx \frac{4\int w^3(r)\phi^4(r)\left|1+\frac{1}{\phi(r)P(k_{ln})}\right|^2 j_l^4(k_{ln}r)\,\mathrm{d}^3r}{\int w^4(r)\phi^4(r)\left|1+\frac{1}{\phi(r)P(k_{ln})}\right|^2 j_l^4(k_{ln}r)\,\mathrm{d}^3r} + \frac{4\int w(r)\phi^2(r)j_l^2(k_{ln}r)\,\mathrm{d}^3r}{\int w^2(r)\phi^2(r)j_l^2(k_{ln}r)\mathrm{d}^3r} = 0, \tag{4.70}$$

with a solution

$$w(r) \approx \frac{1}{1+P(k_{ln})\phi(r)}, \tag{4.71}$$

where an unimportant amplitude factor has been ignored. Feldman, Kaiser & Peacock (1993) also derived this expression in a similar, if slightly less general way; both derivations make use of the properties of the Gaussianity of the density field.

In addition to reducing the measured rms errors of the statistics, another advantage of this weighting scheme is that it makes the window functions (mixing matrices $\Phi_{ll'nn'}^{mm'}$) narrower in k-space than for equal weighting. In particular, it reduces the risk of contamination by nonlinear modes. A more general analysis, which can further narrow the window functions at the expense of increasing the error bars, can be found in Tegmark (1995).

HT extends this analysis to the much more difficult $\beta \neq 0$ case, although this is not actually used.

4.4 Discussion

This spherical harmonic formalism seems to be the ideal method for analysing recent all-sky surveys. The idea has been around for a long time (Peebles 1973), but perhaps only the almost complete sky coverage of the *IRAS* survey made its benefits transparent. With the inclusion of spherical Bessel functions to create a 3-dimensional transform it becomes the obvious method for all-sky redshift surveys, retaining the power spectrum advantage of a clear division between linear and nonlinear regimes. However, even for partial sky coverage, it would seem to be the ideal formalism because of the radial nature of the distortion and the natural separation between angular and radial selection. It appears almost inevitable that observations radiating out from a single vantage point in the universe will be analysed more naturally in spherical geometry.

Chapter 5

SPHERICAL HARMONIC ANALYSIS – APPLICATIONS

The very powerful method developed in Heavens & Taylor (1995; hereafter HT) and chapter 4, is here tested on simulations and applied to appropriate datasets, near-all-sky surveys with a spherical geometry. An approach is used which allows a non-parametric fitting of the shape and amplitude of the large-scale *real-space* power spectrum simultaneously with β. The only other estimate of the real-space power spectrum from an *IRAS* redshift survey is the cross-correlation analysis of Saunders, Rowan-Robinson & Lawrence (1992), which made small-angle approximations. Their small scale analysis complements the large-λ work presented here.

The redshift surveys used are the 1.2Jy (Fisher *et al.* 1995) survey and the PSCz (Saunders *et al.* 1995), both drawn from the near-all-sky *IRAS* Point Source Catalogue. The work detailed is this chapter is based on that contained in Ballinger, Heavens & Taylor (1995), and Tadros *et al.* (in preparation).

5.1 Parameter Estimation Approach

As discussed in chapter 4, for underlying Gaussian fluctuations in the linear regime, the observed quantities D_μ will have a multivariate Gaussian distribution. We can therefore construct a Gaussian likelihood function for the parameter β and the (parametrised) power spectrum, which appears in the covariance matrix $C_{\mu\nu} \equiv \langle D_\mu D_\nu \rangle$ (see chapter 4):

$$\mathcal{L}[D|\beta, P(k_{ln})] = (\det \mathbf{C})^{-1/2} \exp\left(-\frac{1}{2}\sum_{\mu\nu} D_\mu \mathbf{C}^{-1}_{\mu\nu} D_\nu\right) \tag{5.1}$$

The mean values in this model are fixed, the covariance matrix depends on the parameter β and the function $P(k)$. Equation (4.59) can be expressed in terms of "A-matrices", coefficients of the various powers of β,

$$\mathbf{C} = (\mathbf{A}_0 + \beta \mathbf{A}_1 + \beta^2 \mathbf{A}_2) P(k) + \boldsymbol{\Lambda}. \tag{5.2}$$

When exploring a parameter space using a likelihood technique, it is important that the A-matrices can be recalculated quickly at each point. Two very effective methods for this are presented in this chapter.

There are a number of ways to parametrise the power spectrum:

1. Fix the shape of the power spectrum but allow the amplitude to be free, as in HT.
2. Fit the parameters of a standard parametric power spectrum form.
3. Use the *Stepwise Maximum Likelihood* technique to fit the shape of the power spectrum in an almost non-parametric way.

The first technique was adopted by HT for the 1.2Jy survey. As the parameter space is two dimensional the likelihood contours can be mapped in detail. The three sub-matrices of the covariance matrix (see equation (5.2)) need only be calculated once and are then just multiplied by the appropriate factor at each point in parameter space. This does, however, assume a real-space power spectrum form which will have been fitted from another, possibly less reliable, technique. It would seem better to derive the shape of the real-space power spectrum as well in a fully consistent analysis.

The second method is less restrictive but forces the power spectrum to correspond to whichever model has been picked, *e.g.* the two power law fit of Peacock (1992) or the CDM model of Bardeen *et al.* (1986). It may also slow the analysis considerably, as the the sub-matrices of equation (5.2) will have to be calculated each time. The third method, the stepwise maximum likelihood technique, solves both of these problems.

5.1.1 Stepwise Maximum Likelihood

This method fits the shape of a function in an essentially model-independent way – see Efstathiou, Ellis & Peterson (1988). The function is divided into bins, and the amplitude of these bins is allowed to vary.

We exploit the fact that we have a good estimate for the power spectrum from previous studies, $P_0(k)$, so we write

$$P(k) = P_0(k) \times 10^{f_i}. \tag{5.3}$$

where the original shape is multiplied by the free function 10^{f_i}. The subscript $i = 1, 2, 3, \ldots$ labels each wavenumber bin corresponding to the range $k_i^{\min} < k < k_i^{\max}$. The parameters are introduced in this exponential form to ensure that the power spectrum is positive.

The parameters for the maximum likelihood fit are the distortion parameter β and the coefficients f_i; $i = 1, \ldots, n$. This construction essentially amounts to specifying the power spectrum at a number of wavenumbers, but the form chosen allows us to start the maximum likelihood search conveniently at $f_i = 0$, since we have a fairly good idea of what to expect. Note that, although the form suggests that the break scale is fixed, adjustment of the f_i can alter this in essentially an arbitrary way, which is the main reason for parametrising the power spectrum in bins. In HT, and in Fisher, Scharf & Lahav (1994), the power spectrum was parametrised to have a specific shape, whereas here it has a free form (apart from bin discreteness).

The covariance matrix of equation (5.2) can be constructed with little computational cost as each submatrix is a linear combination of contributions from each power-

spectrum bin;

$$\mathbf{A}_0 = 10^{f_1}\mathbf{A}_{01} + 10^{f_2}\mathbf{A}_{02} + \ldots, \tag{5.4}$$

where $\mathbf{A}_{01}$ is the contribution to $\mathbf{A}_0$ from the first bin *etc.*. The matrices $\mathbf{A}_{0i}$, $\mathbf{A}_{1i}$ and $\mathbf{A}_{2i}$ need only be constructed once and then the full A-matrices can be created each time from a simple linear sum. This requires far less computer time than method 2, fitting for a standard power spectrum parameter such as Γ. That would require the A-matrices to be constructed afresh from window function convolutions each time a parameter changed.

The maximizing procedure is straightforward, the $(n+1)$D parameter space is searched until a maximum of the likelihood function $\mathcal{L}[D|\beta, f_i; i = 1, \ldots, n]$ is found. This generalizes the method of HT, where, in effect, the f_i were all constrained to move up or down together.

If n is high, exploring the likelihood space with a grid can become very computationally expensive. However, the maximum likelihood solution is easily found, since we have a good starting guess for the parameters, and can use a standard library routine to maximise the Likelihood function with respect to the parameters. Ideally, we would like to construct complete hypersurfaces of constant likelihood, but this is impractical because of the dimensionality of the problem. Errors can be obtained by searching along the parameter axes until the likelihood drops by $\sqrt{\mathrm{e}}$, but note that this ignores the correlations between parameters. This means that the errors are *conditional*: for each parameter they are calculated assuming that the other parameter values are known exactly – see chapter 6 and Press *et al.* (1992; §15) for a discussion. Note in particular that there is a strong covariance between the amplitude of the power spectrum and β, and these errors do not properly reflect this (see HT; Fisher, Scharf & Lahav 1994). Clearly a calculation of the *marginal* errors – accounting for the correlations – would be preferable; this would, however, require a detailed exploration of the high dimensional parameter space, a daunting computational task.

Note that by parametrising in bins, the fitted power spectrum is discontinuous. In principle there might be a case for adjusting the shape of the power spectrum so that it passes continuously through the maximum likelihood points; in the event there is little

to be achieved by this, since the error bars are large enough so that a smooth curve will be a perfectly acceptable fit, and it would no longer be possible to rapidly produce a covariance matrix using equation (5.4).

5.2 Tests on Simulated Surveys

The introduction to this section and some of the first subsection are based more directly on Tadros *et al.* (1997; in preparation)

The methods described above and in chapter 4 have been tested with mock PSCz redshift surveys drawn from a range of N-body simulations.

5.2.1 Numerical simulations

The dissipationless N-body simulations consist of 3 ensembles of 9 simulations each. The first two ensembles are of CDM-like universes and contain 160^3 particles within a periodic computational box of length $\ell_b = 600\,h^{-1}$ Mpc. These simulations are similar to the ones used and discussed by Croft & Efstathiou (1994a) but employ more particles within a larger computation box. (The simulations of Croft and Efstathiou use 100^3 particles within a box of size $\ell_b = 300\,h^{-1}$ Mpc). The simulations were run with the particle-particle-particle-mesh (P^3M) code described by Efstathiou *et al.* (1985) and model the gravitational clustering of dark matter in a CDM dominated universe with scale invariant initial density fluctuations (see chapter 1). These two ensembles are as follows: the standard CDM model (Davis *et al.* 1985), *i.e.* a spatially flat universe with $\Omega_0 = 1$ and $h = 0.5$ (the SCDM ensemble); a spatially flat low density CDM universe with $\Omega_0 = 0.2$, $h = 1.0$, and a cosmological constant contribution $\Omega_\Lambda = \Lambda/(3H_0^2) = (1 - \Omega_0) = 0.8$ (the LCDM ensemble). The initial power spectra of the models are generated from the fitting function

$$P(k) \propto \frac{k}{\left[1 + \left(ak + (bk)^{\frac{3}{2}} + (ck)^2\right)^\nu\right]^{\frac{2}{\nu}}}, \tag{5.5}$$

where $\nu = 1.13$, $a = (6.4/\Gamma)\,h^{-1}\,\mathrm{Mpc}$, $b = (3.0/\Gamma)\,h^{-1}\,\mathrm{Mpc}$ and $c = (1.7/\Gamma)\,h^{-1}\,\mathrm{Mpc}$. Equation (5.5) is a good approximation to the linear power spectrum of scale-invariant CDM models with low baryon density, $\Omega_b \ll \Omega_0$ (Bond & Efstathiou 1984). The parameter Γ in equation (5.5) is equal to 0.5 for the SCDM ensemble and $\Gamma = 0.2$ for the LCDM ensemble. For pure CDM-type models the shape parameter is given by $\Gamma = \Omega h$. If the model contains baryonic matter then it is claimed that it can be rescaled so $\Gamma = \Omega h e^{-\Omega_B/2}$, see Sugiyama & Sato 1995 and Peacock & Dodds (1994), but this has been questioned in detail – Eisenstein & Hu (1997).

The third ensemble of 9 simulations is a spatially flat mixed dark matter model (MDM) in which CDM contributes $\Omega_{CDM} = 0.6$, baryons contribute $\Omega_b = 0.1$ and massive neutrinos contribute $\Omega_\nu = 0.3$. This set of simulations was run in a box of length $300\,h^{-1}\,\mathrm{Mpc}$ with 100^3 particles. The initial conditions for these simulations are generated from the power spectrum given in equation (1) of Klypin *et al.* (1993). In the MDM models, the thermal motions of the neutrinos are ignored and so they follow the evolution of a collisionless cold component with $\Omega_0 = 1$. The MDM simulations are exactly those used by Croft & Efstathiou (1994b).

In the analysis presented below, simulated data is used to model a survey out to a (comoving) distance of at least $240\,h^{-1}\,\mathrm{Mpc}$, this means that the MDM models are not large enough to properly simulate the survey without a significant amount of "wrapping" of the simulation box. The larger CDM-like models however are superior in this respect and no wrapping is required.

The final output times of the models are chosen to approximately match the microwave background anisotropies measured in the first year COBE maps (Smoot*et al.* 1992) ignoring any contribution from gravitational waves; $\Delta(k) = 180k^2T(k)$. Thus the *rms* mass fluctuations in spheres of radius $8\,h^{-1}\,\mathrm{Mpc}$ are $\sigma_8 = 1$ for the CDM-like models and $\sigma_8 = 0.67$ for the MDM model. The amplitude of the temperature anisotropies measured from the four year COBE maps (Wright *et al.* 1996) is slightly higher than from the first year results. However, for the purposes of testing the techniques described above, this discrepancy is unimportant.

5.2.2 Construction of mock PSCz surveys

The mock PSCz redshift surveys (created by G. Efstathiou) are drawn from the N-body simulation distributions as follows:

1. An observer, situated at an arbitrary position in the simulation box, carves out a sphere of radius $r_{\rm max} = 240\,h^{-1}\,{\rm Mpc}$ (for the MDM simulations this requires some wrapping of the N-body box).

2. Mass points are selected according the the PSCz selection function:

$$\phi(\Delta) = \phi_\star \frac{10^{(1-\alpha)\Delta}}{(1+10^{\gamma\Delta})^{\frac{\eta}{\gamma}}}, \tag{5.6}$$

   where the values of the parameters in equation (5.6) are $\phi_\star = 0.006794(\,h^{-1}\,{\rm Mpc})^{-3}$, $r_0 = 1.95113\,h^{-1}\,{\rm Mpc}$, $\alpha = 1.88654$, $\eta = 4.38595$ and $\gamma = 1.54554$. These parameters were derived using the methods described by Mann, Saunders & Taylor (1996).

3. Mass points are moved to their redshift space positions (**s**) via the equation

$$\mathbf{s} = \mathbf{r}\left(1 + \frac{\mathbf{v}\cdot\hat{\mathbf{r}}}{H_0}\right) \tag{5.7}$$

   where H_0 is the Hubble constant and we have full information on the velocity **v** of each particle.

4. Cartesian x, y, z coordinates are converted to galactic latitude and longitude and the PSCz sky-mask (see section 5.3.2) is applied to the distribution.

5. Mass points are assigned a flux in Janskies according to the selection function.

6. The PSCz sky mask is applied to the resulting catalogue.

This procedure produces mock PSCz surveys with approximately the same density (and therefore the same number) of mass points as there are galaxies in the PSCz survey to a proper distance of $r_{\rm max}\,h^{-1}\,{\rm Mpc}$. Mass points are taken to represent galaxies throughout the analysis, so the linear bias parameter is unity. We have made no attempt to choose 'galaxies' from the simulation density distribution as this introduces

extra complications and is unnecessary for the statistical testing purposes for which we employ the mock catalogues.

5.2.3 Simulation Results – Fixed $P(k)$

Figures 5.1, 5.2 and 5.3 show the result of applying the two parameter analysis to the mock PSCz catalogues. In all cases, the boundary was fixed at $r_{\rm max} = 240\,h^{-1}\,{\rm Mpc}$, the wavenumber cutoff at $k = 0.1\ h\,{\rm Mpc}^{-1}$ (as used by HT). The maximum l and n values are 21 and 8. For the convolutions we have included modes up to $l = 30$, $n = 20$. This is identical to the main PSCz analysis of section (5.3.2).

For the calculation of the scattering matrix (equation 4.33), the three-dimensional velocity dispersion σ was measured from the simulations for each model. We have used the values: $\sigma = 1160{\rm km\,s^{-1}}$, $587{\rm km\,s^{-1}}$ and $859{\rm km\,s^{-1}}$ for the SCDM, LCDM and MDM models respectively. This is not ideal, these values may not correspond exactly to the "effective σ" of equation (4.33) – see chapter 2. It does, however, seem to be a good approximation in practice, even for SCDM where the correction is large. In the analysis of the CDM-like mock catalogues we have assumed a real-space power spectrum of the form given in equation (5.5). For the MDM model, a power spectrum of the form of equation (1) of Klypin *et al.* 1993 was used.

Figures 5.1, 5.2 and 5.3 show contours of likelihood in the $\beta - \Delta_{0.1}$ plane for the SCDM, LCDM and MDM mock catalogues respectively. The contours are plotted at intervals of 0.5 in log–likelihood and the x-axis is labeled at intervals of 0.1 in $\Delta_{0.1}$. The maximum likelihood values of β and $\Delta_{0.1}$ are indicated by a cross. In the limit that the distribution of the parameter values is a bivariate Gaussian, this contour spacing has very convenient statistical properties. For example the projection of the innermost contour onto a parameter axis gives a 1σ marginal error: $\sim 68\%$ confidence interval – see Press *et al.* (1992) and chapter 6. In the approximation that the distribution of parameters is a bivariate Gaussian, the sixth contour ($\Delta\chi^2 = 3$) encloses a $\sim 95\%$ confidence region, see Press *et al.* (1992; §15). Thus about one figure in 20 should have the true value lying outside this final contour. In these examples, 2 realisations out of

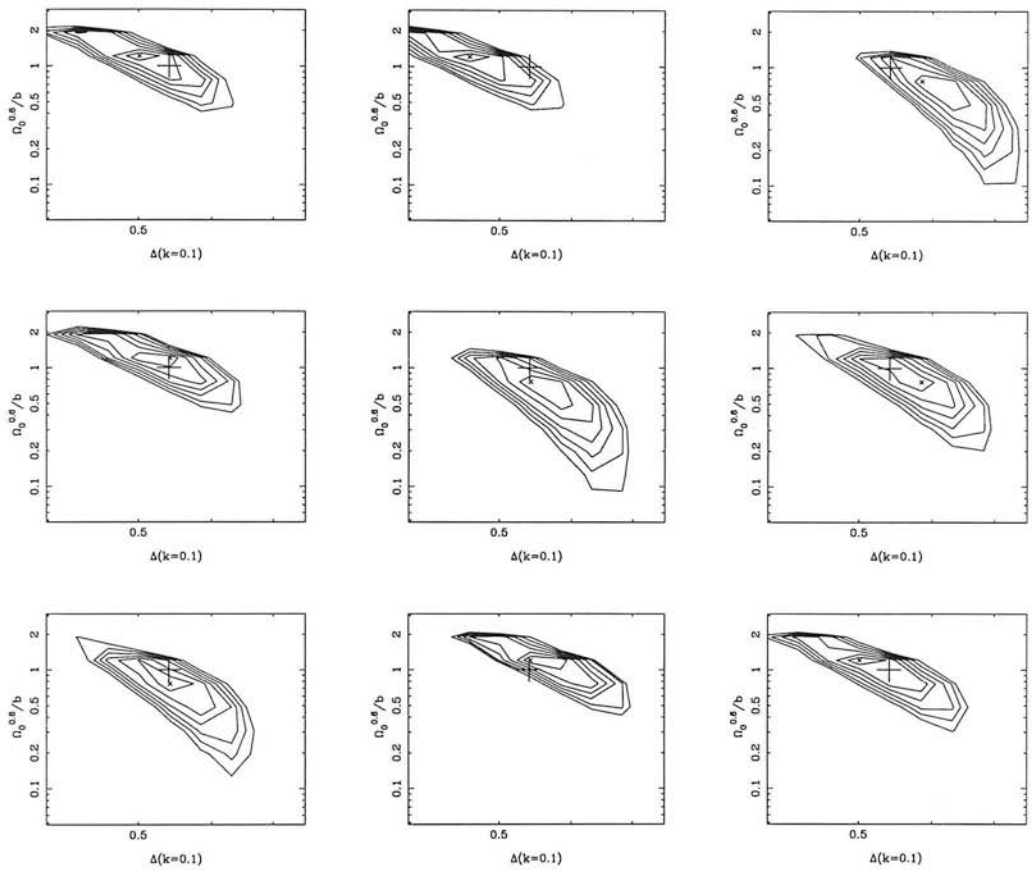

Figure 5.1: Figure shows contours of likelihood in the β - $\Delta(k = 0.1)$ plane for 9 realizations of the PSCz redshift survey drawn from SCDM N-body simulations. The contours are plotted at intervals of $\delta \ln \mathcal{L} = 0.5$. The x axis is plotted in steps of 0.1 in $\Delta(k = 0.1)$. The vertical cross marks the true parameter values of $\Delta_{\rm true} = 0.54$, $\beta_{\rm true} = 1$.

27 are like this (the first and sixth of Fig 5.3), roughly the number to be expected.

The statistical model can be tested by comparing two different measures of the expected parameter errors for a single realisation. The formal "internal" errors on the values of β and $\Delta(k = 0.1)$ are given by the projection of $\delta \ln \mathcal{L} = -0.5$ onto the parameter axes, *i.e.* the projection of the first plotted contour. The mean value of the internal error is found by averaging over the internal errors of the 9 realisations. This is the error expected from the model, although it is not strictly correct because the model parameter error depends on the actual value of the maximum likelihood (ML)

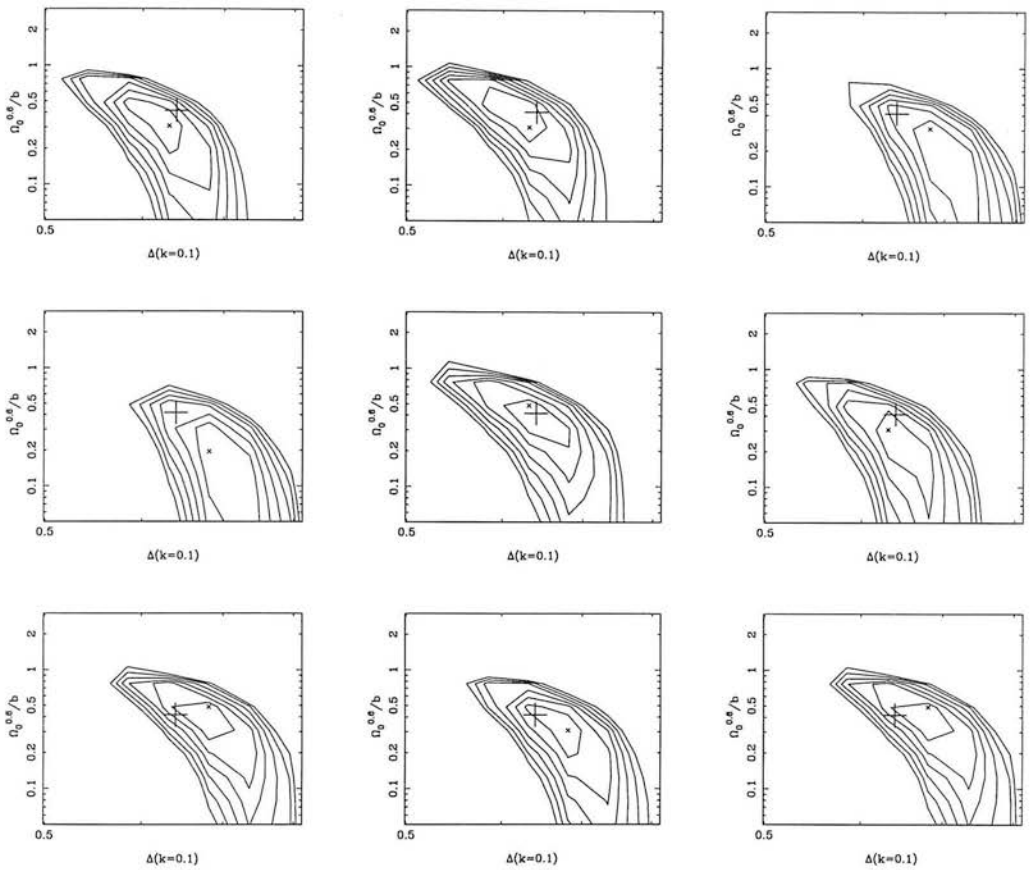

Figure 5.2: Contours of likelihood in the β - $\Delta(k = 0.1)$ plane for 9 realizations of the PSCz redshift survey drawn from LCDM N-body simulations. Again, contours are plotted at intervals of $\delta \ln \mathcal{L} = 0.5$ and the x axis is plotted in steps of 0.1 in $\Delta(k = 0.1)$. The vertical cross marks the true parameter values of $\Delta_{\rm true} = 0.64$, $\beta_{\rm true} = 0.38$.

parameter. Alternatively, the error can be estimated from the scatter of the ML point between realisations, *i.e.* the standard deviation of the observed distribution of ML parameter values. This second error bar provides us with an "external" estimate of the error. In addition, the scatter of the internal error gives an estimate of the 1-σ "error on the error", which can be used to see if it is consistent with the external value.

The internal and external errors are plotted in Table (5.2. The internal error measured from the projection of the first contour level provides a reasonable estimate of the true error, as can be seen from the fact that the external and internal error estimates are in good agreement. The only case where this is not true is in the MDM model

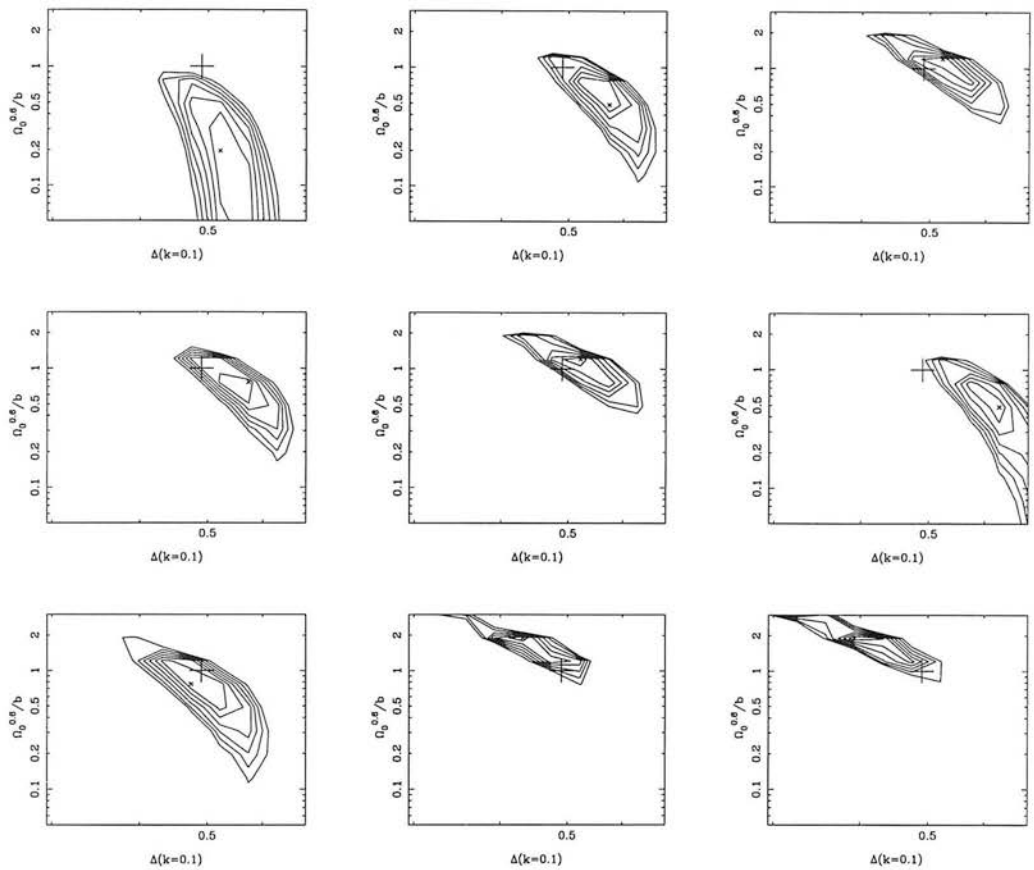

Figure 5.3: Contours of likelihood in the β - $\Delta(k = 0.1)$ plane for 9 realizations of the PSCz redshift survey drawn from MDM N-body simulations. The vertical cross marks the true parameter values of $\Delta_{\rm true} = 0.49$, $\beta_{\rm true} = 1$.

where the external error appears to be too large. This can possibly be explained by the fact that these simulations are not large enough to model the PSCz redshift survey out to a proper distance of $r_{\rm max} = 240\,h^{-1}\,{\rm Mpc}$.

The average recovered value of β and $\Delta(k = 0.1)$ over the nine realizations of the PSCz survey for each of the cosmological models, together with the true values of these parameters are presented in Table 5.1. The errors are estimated by dividing the internal errors by 3 ($\sqrt{9}$). The recovered values of β and $\Delta(k = 0.1)$, are in very good agreement with the true values, well within the expected 1-σ errors. The values for β seem perhaps too accurate – the mean error may be overestimated because of the large variation in ML β values with different individual errors. This seems unlikely, however,

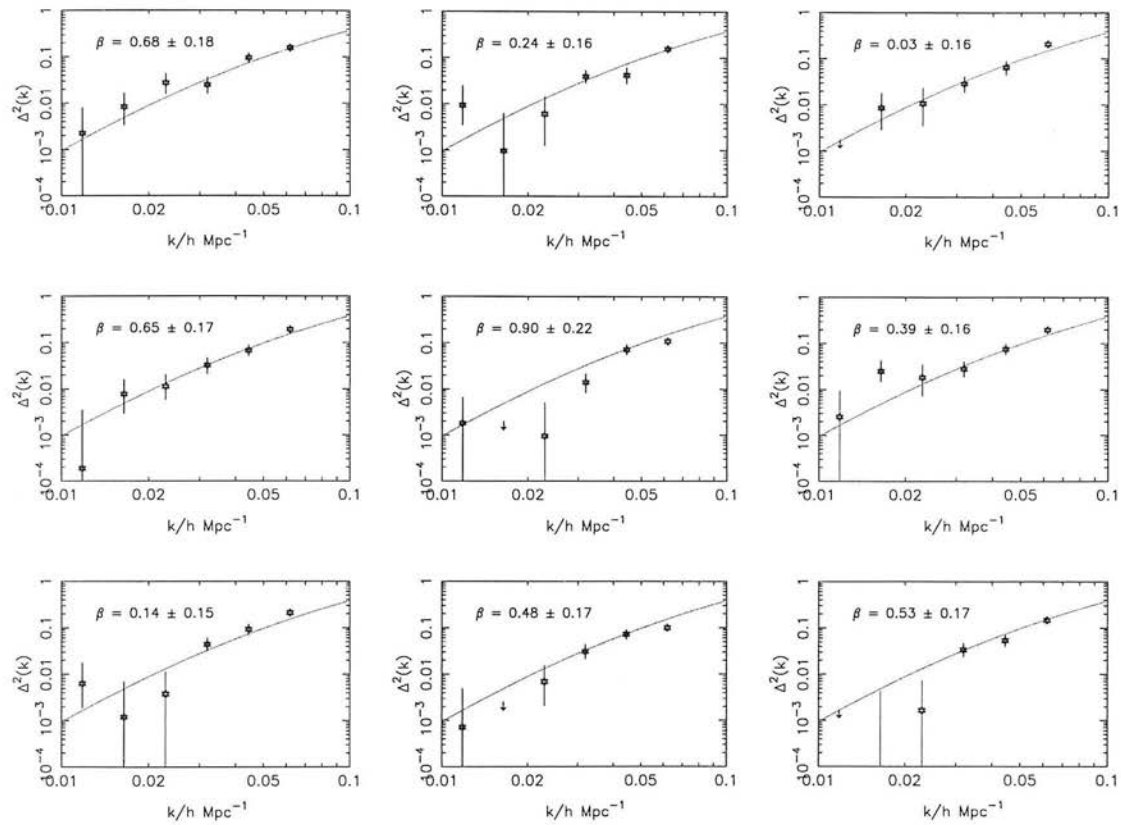

Figure 5.4: Figure shows the stepwise power spectrum estimation for the nine LCDM simulations with $l_{\max} = 19$, $n_{\max} = 7$, $r_{\max} = 300\,h^{-1}\,\mathrm{Mpc}$ and $k_{\max} = 0.073\,h\,\mathrm{Mpc}^{-1}$. The corresponding estimate of β is given alongside each curve.

given that the internal and external errors match up quite well for SCDM and LCDM.

It is remarkable that the correction applied here for the non-linear effects introduced by the small-scale velocity field is adequate even for the SCDM model, which has a very high three-dimensional velocity dispersion ($\sigma = 1160\mathrm{km\ s}^{-1}$), especially considering the objections raised in chapter 2. Without this correction the recovered value of β would be significantly lower than the true value as non-linear effects become important even for relatively small wavenumbers.

From these tests on simulated PSCz redshift surveys we can be confident that the analysis described in this paper provides an unbiased estimate of the parameters β and $\Delta(k = 0.1)$. Furthermore, the errors obtained by examining the likelihood contours

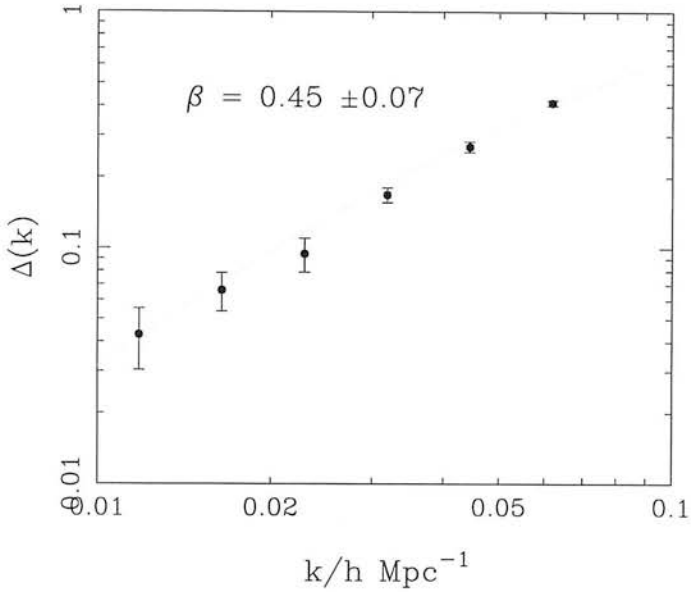

Figure 5.5: The LCDM power spectrum estimates averaged together, showing good agreement with true power spectrum, and the correct $\beta_{LCDM} = 0.38$.

provide a reasonable estimate of the true error on β and $\Delta(k = 0.1)$.

Model	$\beta_{\rm rec}$	$\beta_{\rm true}$	$\Delta_{\rm rec}$	$\Delta_{\rm true}$
SCDM	1.01 ± 0.06	1.00	0.53 ± 0.01	0.54
LCDM	0.36 ± 0.05	0.38	0.66 ± 0.01	0.64
MDM	1.00 ± 0.07	1.00	0.51 ± 0.02	0.49

Table 5.1: Results of applying the two-parameter spherical-harmonic model to mock PSCz redshift surveys drawn from three cosmological models. Columns 1 and 3 give the recovered values of β and $\Delta(k = 0.1)$, averaged over nine realizations. These are to be compared with the true values of the parameters, given in columns 2 and 4.

5.2.4 Simulation Results – Fitting $P(k)$

The testing of the stepwise method has been somewhat more limited but looks very promising. The LCDM simulations mentioned above were used, but without an angular mask. The transform used $l_{\rm max} = 19$, $n_{\rm max} = 7$, $r_{\rm max} = 300\,h^{-1}\,$Mpc and $k_{\rm max} = 0.073\,h\,{\rm Mpc}^{-1}$. The power spectrum was divided into six bins in the range $0.01 < k < 0.073$, equally spaced in $\log k$. The lowest k bin was extended to include underlying $k < 0.01\,h\,{\rm Mpc}^{-1}$, the upper to include $k > k_{\rm max}$. All of the observed modes analysed

Model	$\text{Error}_\beta^{\text{ext}}$	$\text{Error}_\beta^{\text{int}}$	$\text{Error}_\Delta^{\text{ext}}$	$\text{Error}_\Delta^{\text{int}}$
SCDM	0.24	0.17 ± 0.05	0.03	0.04 ± 0.01
LCDM	0.12	0.15 ± 0.04	0.03	0.05 ± 0.01
MDM	0.60	0.20 ± 0.05	0.06	0.03 ± 0.01

Table 5.2: Comparison of internal and external parameter errors from the 3 simulations.

were from within the stated range of k.

The test was performed for seven simulations, using a NAG routine[1] to search through the 7D parameter space and find the maximum of the likelihood function given in equation (4.64). As discussed in section (5.1.1), the error bars were then found by searching along each parameter axis until $\ln \mathcal{L}$ had fallen by 0.5 – the "1σ" error. Note that this only gives conditional errors – correlations between parameters are expected. Results are plotted on top of the true real-space power spectrum of the simulations and corresponding β measurements are given; the expected value of β for these simulations is $\beta_{\text{LCDM}} = 0.38$. The results show excellent agreement with the expected values, as can be seen when they are averaged together in fig 5.5. This appears to be a very effective method for finding the large-scale real-space power, particularly when coupled with the promising error analysis of section (5.2.3).

5.3 Redshift Survey Results

5.3.1 1.2Jy Survey – Non-parametric Power Spectrum

Using the stepwise maximum likelihood technique discussed above, the power spectrum measurements from the *IRAS* 1.2-Jy survey (Fisher *et al.* 1995a) are plotted in Fig. (5.6), in the wavelength range 63–630 h^{-1} Mpc.

Observed modes in the range $0.01 < k < 0.1$ were analysed. The power spectrum was divided into six wavenumber bins, using the method similar to that described in

[1] E04 JAF – a Quasi-Newton search routine, it constantly updates an approximation to the Hessian matrix.

section (5.2.4). The lowest bin has $k < 0.027\ h\ \mathrm{Mpc}^{-1}$, and the next four have a width of $\Delta \log(k) = 0.12$. The final bin has $k > 0.076$. The lowest bin is larger (in $\log k$) than the others because there is so little signal at the large-scale end. Otherwise the analysis follows that of HT; Neumann boundary conditions are imposed at $s = 200\,h^{-1}\,\mathrm{Mpc}$, and, to ensure linearity, all observed modes have $k \leq 0.1\ h\,\mathrm{Mpc}^{-1}$. The monopole ($l = 0$) and dipole ($l = 1$) are not used for reasons discussed in HT. These conditions impose the limits $2 \leq l \leq 17$ and $1 \leq n \leq 6$. Only $m \neq 0$ modes are used for simplicity, giving a total of 302 real and 302 imaginary modes.

Because of the conservative galactic cut, the mask was assumed to be azimuthally symmetric and hence the Likelihood function was split into real and imaginary parts (see sections (4.3.1) and (E.2)). This ignores the additional smaller masked regions, such as the stripes missed by the *IRAS* all sky survey; it was assumed that the effect would be insignificant.

There are a total of seven free parameters here: β and $f_i; i = 1, 6$ (see equation (5.3)). The maximum was found by finding the maximum likelihood along each axis in turn until convergence was attained. The (conditional) error bars are as described in section (5.1.1).

The galaxy sample is exactly that used in the HT analysis. The sample is conservatively cut at $|b| > 10°$ in addition to the 1.2Jy survey mask to try to avoid any residual confusion with galactic sources, leaving 4511 galaxies. The selection function is modelled as

$$\bar{\rho}\phi(r) = ar^{-2\alpha}\left[1 + \left(\frac{r}{r_*}\right)^2\right]^{-\beta_*}, \tag{5.8}$$

where $a = 0.189$, $\alpha = 0.421$, $\beta_* = 1.913$ and $r_* = 50.1$; all units are $h^{-1}\,\mathrm{Mpc}$. Also, the fiducial power-spectrum form used in HT is retained, the two-power-law fit of Peacock & Dodds (1994)

$$P(k) = \frac{2\pi^2(k/k_0)^{1.5}}{k^3[1 + (k/k_c)^{-2.5}]}, \tag{5.9}$$

with $k_c = 0.29\ h\,\mathrm{Mpc}^{-1}$ and the break scale $k_0 = 0.039\ h\,\mathrm{Mpc}^{-1}$. Note again that despite choosing this initial form the stepwise fitting routine allows an almost arbitrary shape for the final power spectrum. The weighting function was also the same as that of HT, with fiducial power $P(k) = 6000h^{-3}\,\mathrm{Mpc}$ in equation (4.71), roughly the power

expected at a wavenumber $k \sim 0.5\, h\,\mathrm{Mpc}^{-1}$.

The parameter β has the likelihood solution

$$\beta = 1.04 \pm 0.3, \tag{5.10}$$

consistent with the result of HT: $\beta = 1.1 \pm 0.3$. In both these cases the errors are conditional – from inspection of Fig. 5 of HT the marginal error is closer to ± 0.5. The first bin has no detection of power (the maximum likelihood solution is $f_1 \to -\infty$), and the upper limit is set by the power for which the likelihood is $1/\sqrt{\mathrm{e}}$ of its maximum value. The final point puts interesting constraints on the power spectrum: keeping all other parameters fixed, the likelihood of a power spectrum with no break extrapolated from small scales (i.e. a featureless $k^{-1.5}$ power law) is ten times less likely than the Peacock & Dodds (1994) fit which turns over gradually to a Harrison-Zeldovich spectrum at a wavenumber of around $0.04h$ Mpc^{-1}. The lines drawn are as follows: the solid line is the extrapolation of a Harrison-Zeldovich spectrum $P(k) \propto k$ from *COBE* measurements, assuming a quadrupole of $20\mu K$; The dotted line is the mixed dark matter model with 13% neutrinos (van Dalen & Schaefer 1992; Taylor & Rowan-Robinson 1992). The dashed line is the two-power law fit of Peacock & Dodds (1994) to a variety of catalogue data, and corresponds to $f_i \equiv 0$ in (5.3). We see that the data are not inconsistent with this fit. All curves assume $\Omega_0 = 1$, $\Omega_\Lambda = 0$. Note that increasing the neutrino fraction in the mixed models increases the discrepancy with the data at the high-k end.

Comparison with Other Determinations

We show in Fig. (5.7) comparisons of our determination of $P(k)$ with other determinations from different surveys. The plot is rather scattered in this long-wavelength regime. One reason is that the QDOT (Feldman *et al.* 1994) and *IRAS* 1.2-Jy (Fisher *et al.* 1993) results are redshift-space power spectra, which are amplified by redshift-space distortions. On small scales where the distant observer approximation holds, the amplification is 28/15 for $\beta = 1$ (Kaiser 1987). Even so, there is some disagreement between the results, the reasons for which are not yet clear, although Efstathiou (1995)

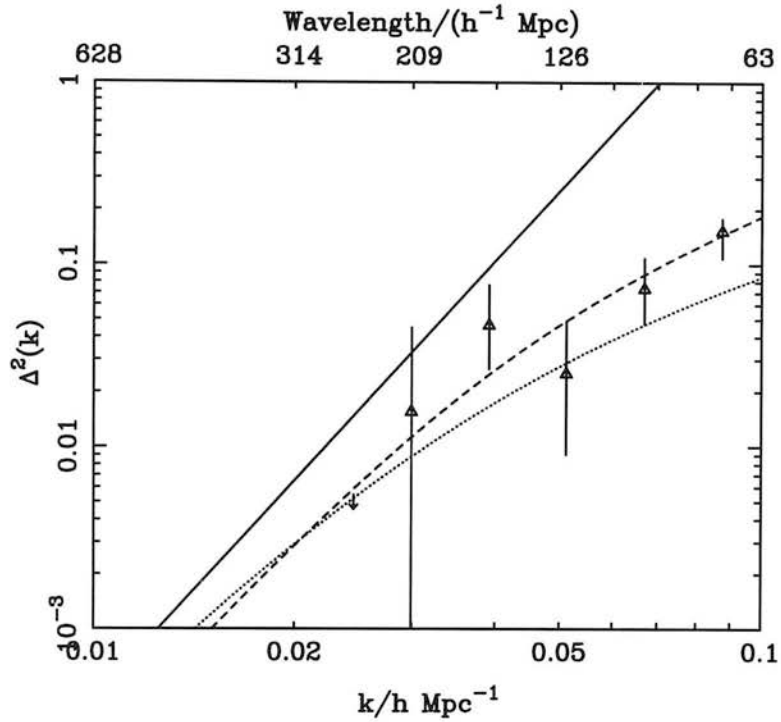

Figure 5.6: The real-space power spectrum of *IRAS* 1.2-Jy galaxies. The solid line is the extrapolation of a Harrison-Zeldovich spectrum $P(k) \propto k$ from *COBE* measurements, assuming a quadrupole of $20\mu K$. The dotted line is mixed dark matter with 13% neutrinos (van Dalen & Schaefer 1992; Taylor & Rowan-Robinson 1992). The dashed line is the two-power law fit of Peacock & Dodds (1994) to a variety of catalogue data. All curves assume $\Omega_0 = 1$.

has suggested that the QDOT result may be sensitive to the sampling in the Hercules cluster. In passing, we also note that the marginally significant bump in $\Delta^2(k)$ at $k \sim 0.04h$ Mpc^{-1} is also found in the redshift-space power spectrum of the combined *IRAS* 1.2Jy and QDOT catalogue (Tadros & Efstathiou 1995). We have not shown the real-space determination of the power spectrum from QDOT of Saunders, Rowan-Robinson & Lawrence (1992), as the small-angle approximations make it unreliable on these scales, but we note that our result matches onto it at $k = 0.1h$ Mpc^{-1}.

We also show the real-space power estimate from deconvolving the projected power in the APM survey (Baugh & Efstathiou 1994), whose statistical errors are clearly much smaller. In the region where our errors are reasonably small ($k \sim 0.05 - 0.1h$

k/h Mpc^{-1}	$\Delta^2(k)$	$\Delta^2(k)$ lower	$\Delta^2(k)$ upper
0.0245	0.0	0.0	0.0053
0.0299	0.016	0.0	0.045
0.0391	0.046	0.027	0.077
0.0511	0.025	0.0091	0.049
0.0669	0.073	0.047	0.11
0.0874	0.15	0.11	0.17

Table 5.3: Power spectrum measurements and limits for the 1.2Jy Survey.

Mpc^{-1}), we see that the *IRAS* 1.2-Jy sample has rather less power than APM. This may arise from differences in the populations of optical and infrared galaxies on much smaller scales. In particular, lower infrared galaxy densities in the cores of rich clusters (*e.g.* Strauss *et al.* 1992) can lead to a bias between the power spectra on very large scales (Mann, Peacock & Heavens 1997).

5.3.2 PSCz Survey

The PSCz is a newer, larger survey selected from the same *IRAS* parent catalogue as the 1.2Jy survey. Across 87% of the sky it is complete down to a flux limit of 0.6Jy – the 1.2Jy sample almost forms a subset of this survey.

A conservative mask was provided to give a highly reliable survey area (Saunders, private communication). The shape of this mask is quite irregular, and is not well approximated by a 'zone of avoidance' at fixed galactic latitude, *i.e.* the mask is not azimuthally symmetric, see Fig 5.8. The **W** matrices can no longer be assumed to be real, and so the off diagonal terms of the covariance matrix, which mix real and imaginary modes, can no longer be assumed to be zero – see section (4.3.1). Inverting the covariance matrix is now a more computationally expensive task, it takes four times longer to invert $2n \times 2n$ matrix than two $n \times n$ ones. Note that this analysis correctly

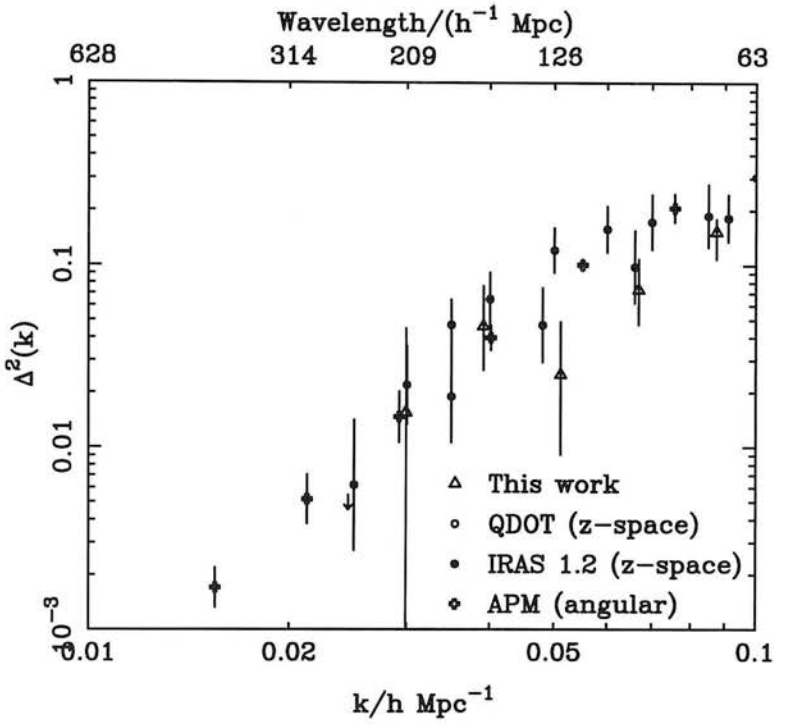

Figure 5.7: The real-space power spectrum of *IRAS* 1.2-Jy galaxies, as in Fig. 1, compared with redshift-space power spectra from the same survey (Fisher *et al.* 1993), and the *IRAS* QDOT survey, also in redshift-space (Feldman *et al.* 1994). The real-space power spectrum deconvolved from the APM catalogue (Baugh & Efstathiou 1994) is also shown.

accounts for all masked regions, including the missing stripes of the *IRAS* survey.

The calculation of the mixing matrices was also improved, correctly accounting for the entirely real $m = 0$ modes which mix in. The effect is small as these modes are not analysed directly, see section (4.3), but they are treated correctly here.

An additional change is an improvement in the weighting scheme. In the PSCz analysis, the galaxies are weighted differently for each mode in the transform, optimising for the expected power spectrum at the corresponding k. This is clearly an improvement over the Feldman *et al.* (1994) and HT techniques, which assume an average power value for the whole transform.

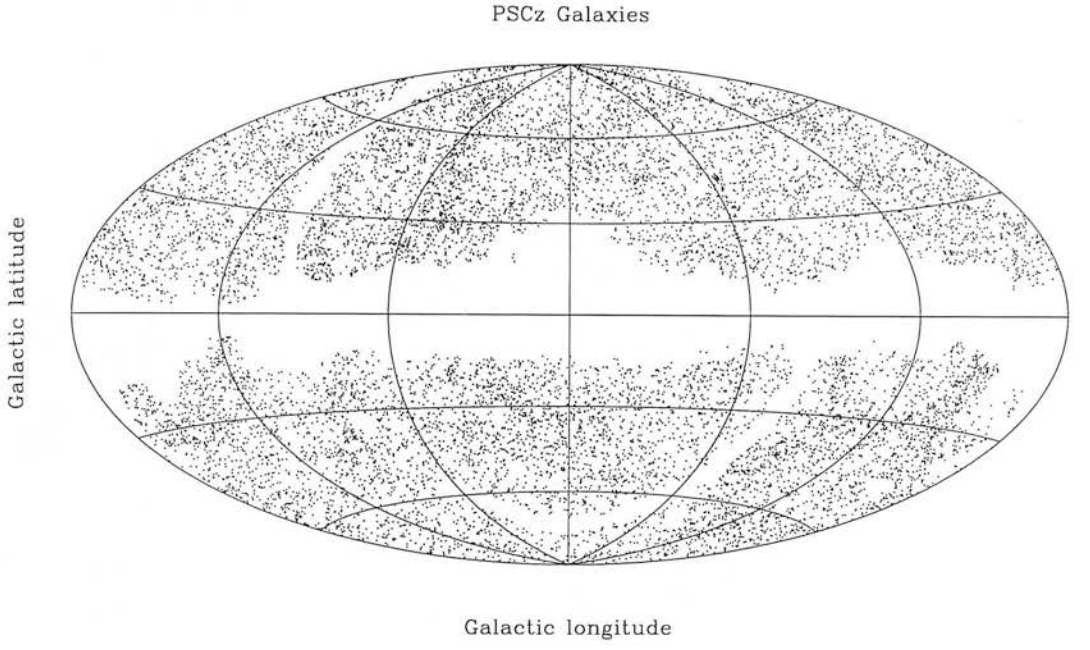

Figure 5.8: The sky distribution of the PSCz sample used in this analysis.

Fixed power-spectrum shape

This section uses the method of HT – only β and the amplitude of the power spectrum vary to give a two parameter model. The small-scale velocity model was used with $\sigma_{\rm p} = 420 {\rm kms}^{-1}$ to suppress any nonlinear effects, although the difference should be very small – see HT.

The two parameter fit was performed with and without the Local-Group correction of equation (4.19) to see if there is any effect. It can be seen from Figs 5.9 and 5.10 that the results are indistinguishable, so the correction can be ignored. This is hardly surprising, the dipole ($l = 1$) term is not analysed directly and only enters via mixing due to the mask.

The Likelihoods are calculated using two different $r_{\rm max}$ values to see if this significantly affects the results.

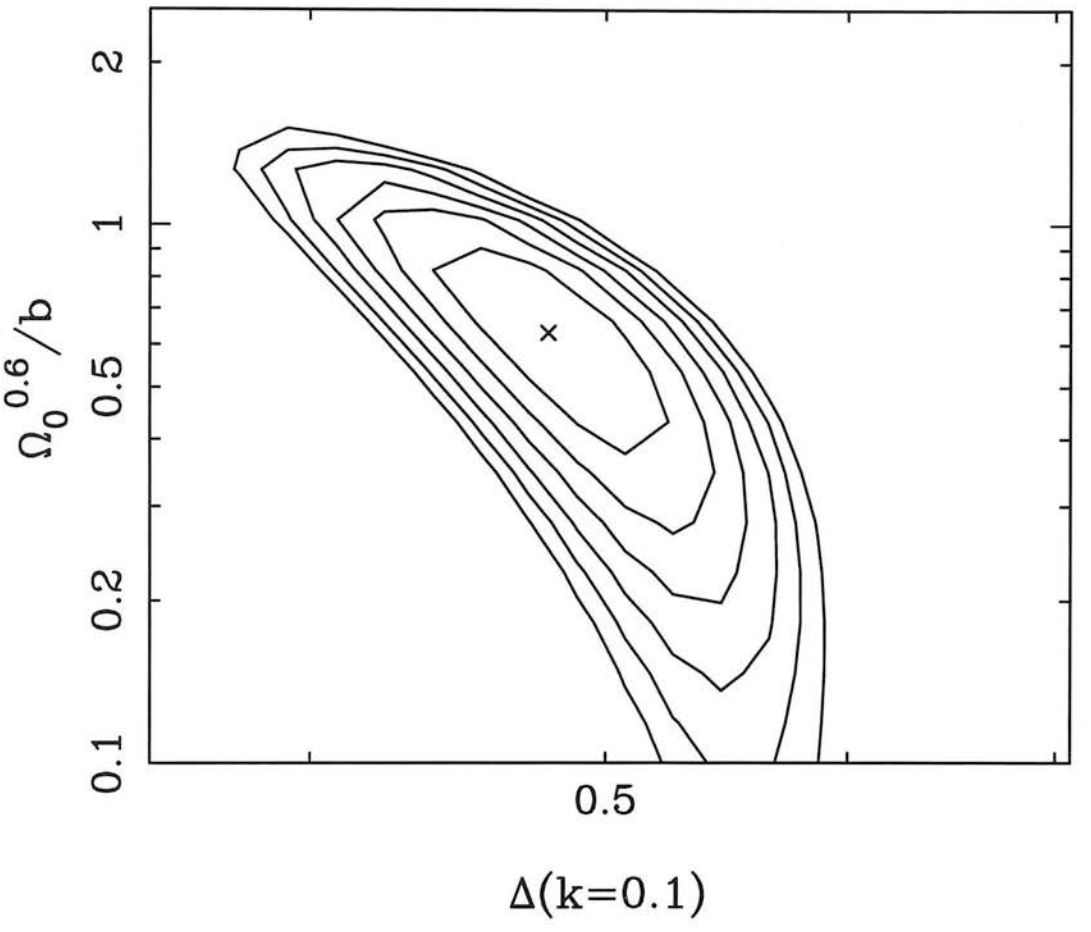

Figure 5.9: Contours of likelihood for β and $\Delta_{0.1}$ for the PSCz survey. The boundary is at $r_{\max} = 240\,h^{-1}\,\mathrm{Mpc}$, with modes analysed up to $k_{\max} = 0.1\ h\,\mathrm{Mpc}^{-1}$, $l_{\max} = 21$ and $n_{\max} = 8$. The contours are separated by intervals of $\Delta \ln \mathcal{L} = 0.5$.

The $1-\sigma$ errors for $r_{\max} = 240\,h^{-1}\,\mathrm{Mpc}$ are

$$\beta = 0.63 \pm 0.25, \tag{5.11}$$

$$\Delta(k = 0.1) = 0.48 \pm 0.03. \tag{5.12}$$

The PSCz catalogue gives a slightly lower value for β than that found for the 1.2Jy survey in equation (5.10) and HT. This new value is more in line with the cartesian multipole results of CFW95. The quoted errors assume that the Likelihoods can be approximated as a bivariate Gaussian – see chapter 6 and Press *et al.* (1992).

The analysis was repeated with a larger value of $r_{\max}$, to see if changing the bound-

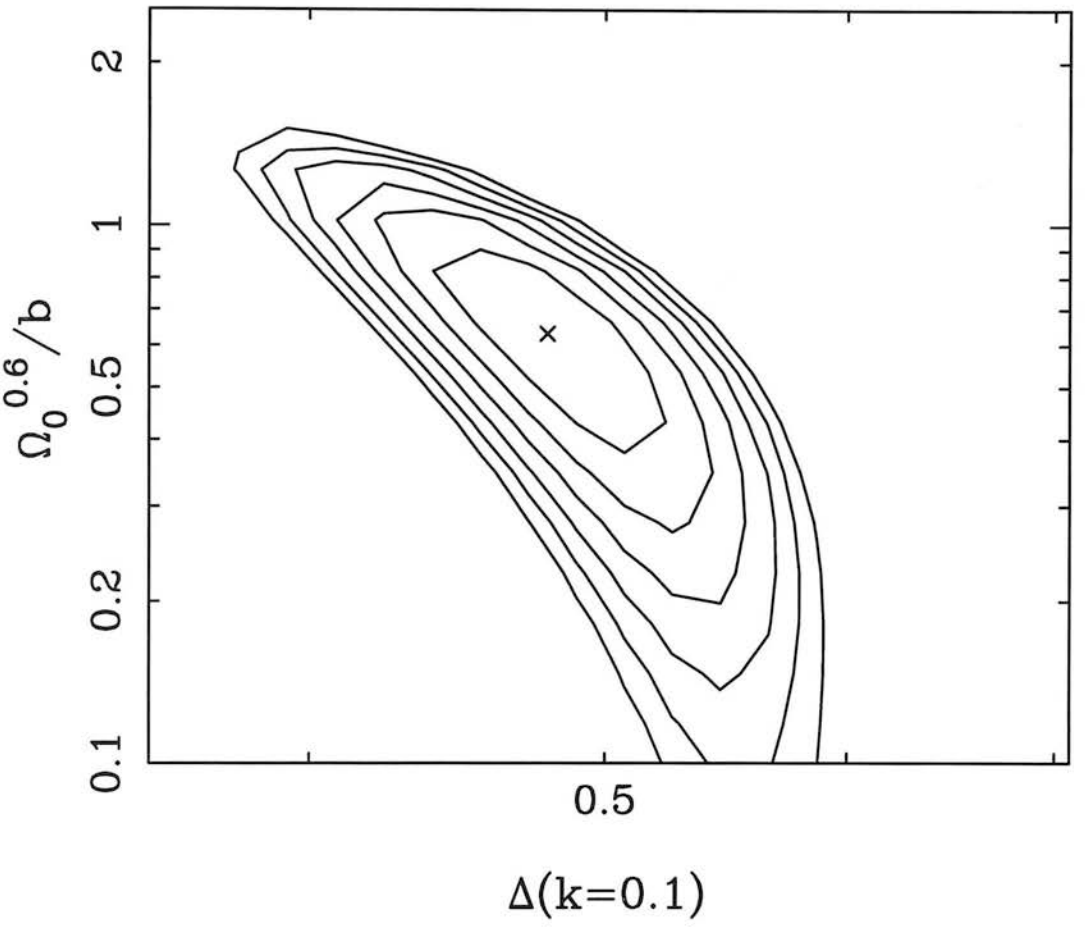

Figure 5.10: As for Fig 5.9, but with the addition of the Local-Group correction of equation (4.19). This correction has a negligible effect.

ary made a significant difference. For the $r_{\max} = 270\,h^{-1}$ Mpc boundary the results are slightly different,

$$\beta = 0.37 \pm 0.3, \tag{5.13}$$

$$\Delta(k = 0.1) = 0.55 \pm 0.05, \tag{5.14}$$

a drop in β and an increase in power. Comparing Figure 5.9 with Figure 5.11, it is clear that the shift of the likelihood ridge itself is moderate, but the maximum contour has slipped along the ridge towards low β. Note again, however, that this does not correspond to the same range of wavenumber as the $240\,h^{-1}$ Mpc analysis, so the results should not be identical.

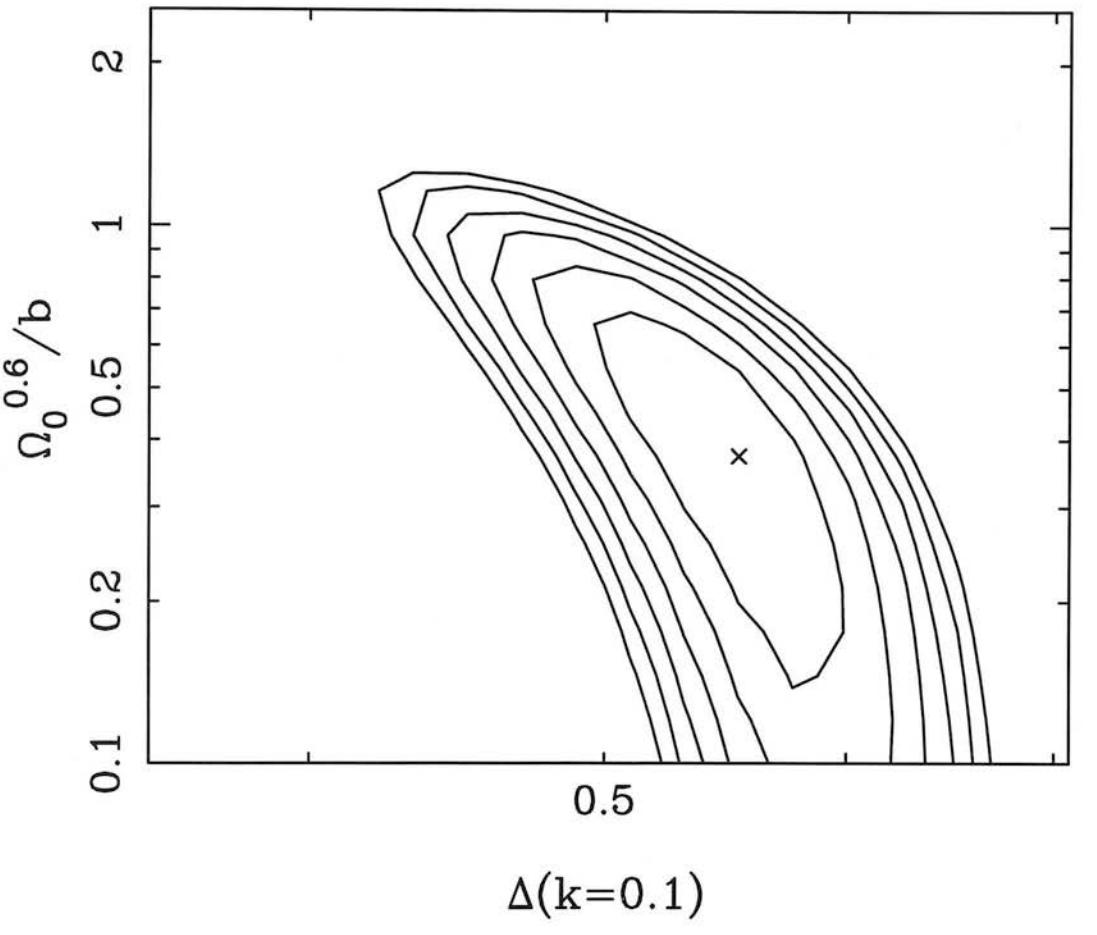

Figure 5.11: Contours of likelihood for β and $\Delta_{0.1}$ for the PSCz survey. The boundary is at $r_{\max} = 270\,h^{-1}\,\mathrm{Mpc}$, with modes analysed up to $k_{\max} = 0.89\,h\,\mathrm{Mpc}^{-1}$, $l_{\max} = 21$ and $n_{\max} = 8$. The contours are separated by intervals of $\Delta \ln \mathcal{L} = 0.5$.

It is possible that this change is a manifestation of the 'Hamilton effect', a decrease in the measured value of β at $r > 80\,h^{-1}\,\mathrm{Mpc}$ which was reported for the QDOT catalogue – see Hamilton (1995). It was suggested that this effect could be caused by a variation of the effective flux limit of ± 0.1Jy on an angular scale of $\sim 7°$ in the *IRAS* Point Source Catalogue from which the redshift surveys are selected. This could lead to a decrease in measured β with increasing r_{max}, although a subtler test would be to change the flux limit of the catalogue, rather than the boundary.

It is clear that there may be a discrepancy in the analyses for different boundary radii. This was tested directly by examining the effect of changing the boundary for a

fixed $k_{\max} = 0.08\, h\,\mathrm{Mpc}^{-1}$, which will keep the number of actual *independent* modes fixed. The results look fairly good – see Fig 5.12. There does seem to be a small difference between $240\, h^{-1}$ Mpc and $270\, h^{-1}$ Mpc, but there is almost no change going up to $300\, h^{-1}$ Mpc. Beyond a certain point, the plots seem to stabilise, but unfortunately the computational expense increases as $r_{\max}^3$. A solution to this problem is offered in the next chapter – data compression.

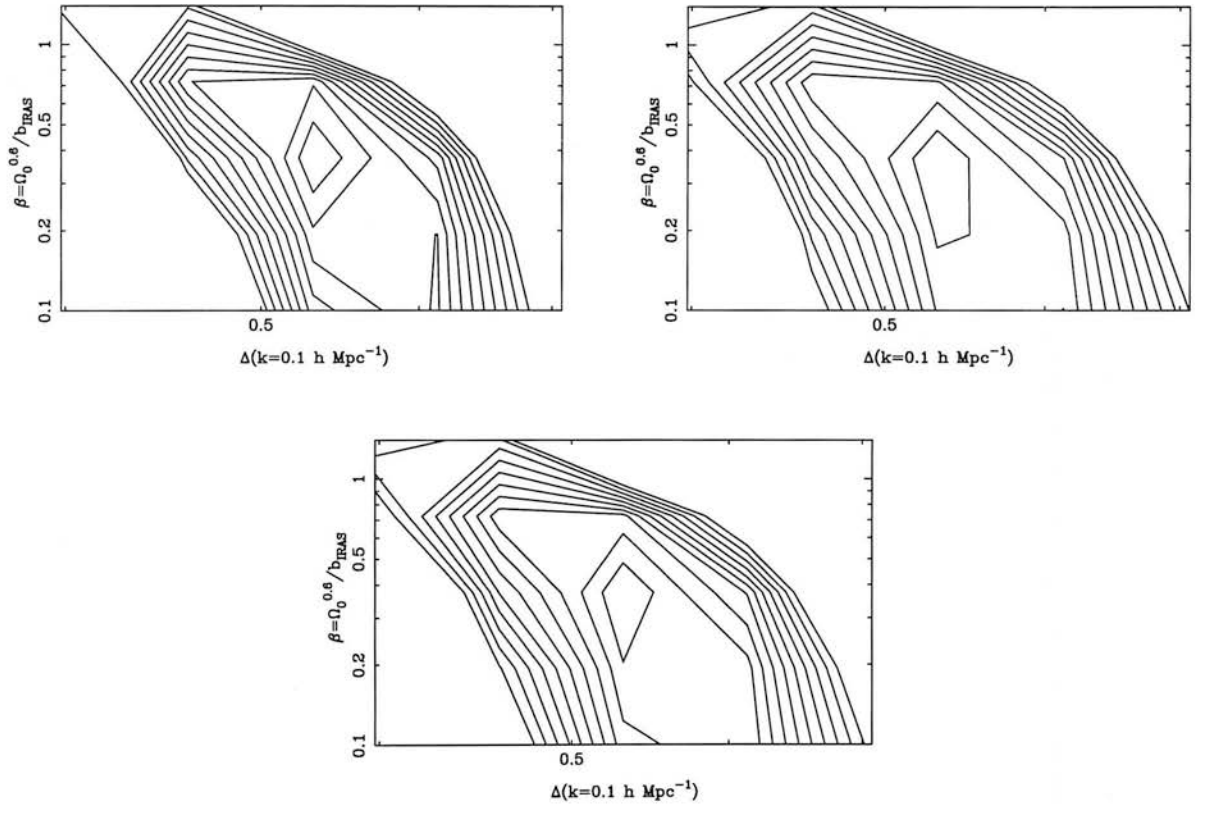

Figure 5.12: Likelihood contours with $k_{\max} = 0.08\, h^{-1}\,\mathrm{Mpc}$ for (a) $r_{\max} = 240\, h^{-1}\,\mathrm{Mpc}$, (b) $r_{\max} = 270\, h^{-1}\,\mathrm{Mpc}$, (c) $r_{\max} = 300\, h^{-1}\,\mathrm{Mpc}$.

Non-parametric Power Spectrum

The stepwise power spectrum was fitted in a similar way to that of the 1.2Jy survey. An additional feature is that the righthand bin is chosen to begin at $k_{\max}$, to show how information beyond the observed modes can be analysed. There is a constraint on this bin due to the effect of mixing, but this is quite weak as can be seen by the large error bar on the final point. Thus there are five bins in the interval $0.01 < k < 0.1\, h\,\mathrm{Mpc}^{-1}$,

each corresponding to an interval of $\Delta \log k = 0.2$. Again, the lowest bin is extended to any $k < 0.01\ h\,\mathrm{Mpc}^{-1}$, but there is an extra "mixing" bin for all modes with $k > 0.1$. As with the 1.2Jy analysis, only modes in the range $0.01 < k < 0.1\ h\,\mathrm{Mpc}^{-1}$ are analysed directly.

Again, reassuringly, the redshift distortion result is similar to that of the two-parameter fit with

$$\beta = 0.61 \pm 0.17, \tag{5.15}$$

although note that this is a conditional error, as are the error bars on the power spectrum.

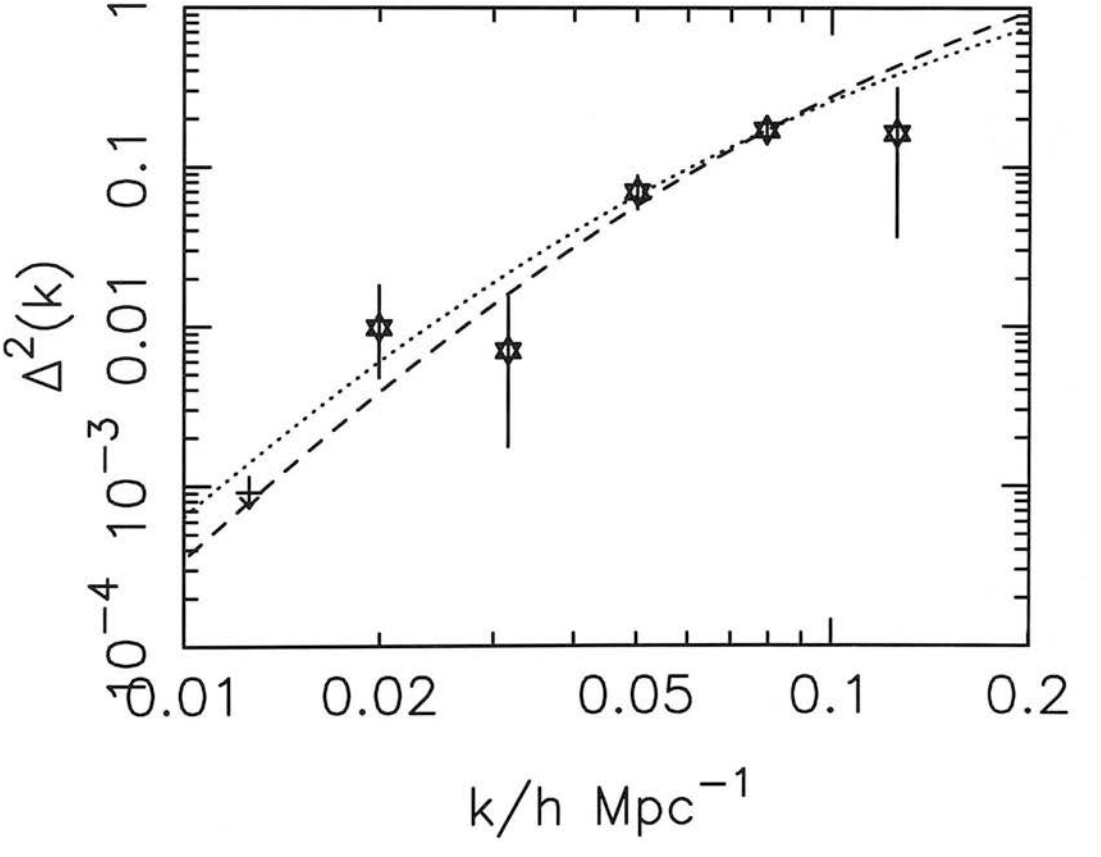

Figure 5.13: The real-space power spectrum of the PSCz redshift survey in dimensionless units. The curves are CDM models with $\Gamma = 0.2$ (dotted), 0.3 (dashed).

The power spectrum again fits low-Γ CDM models very well, although there is some evidence that the break is slightly sharper as suggested by Peacock (1997). It is per-

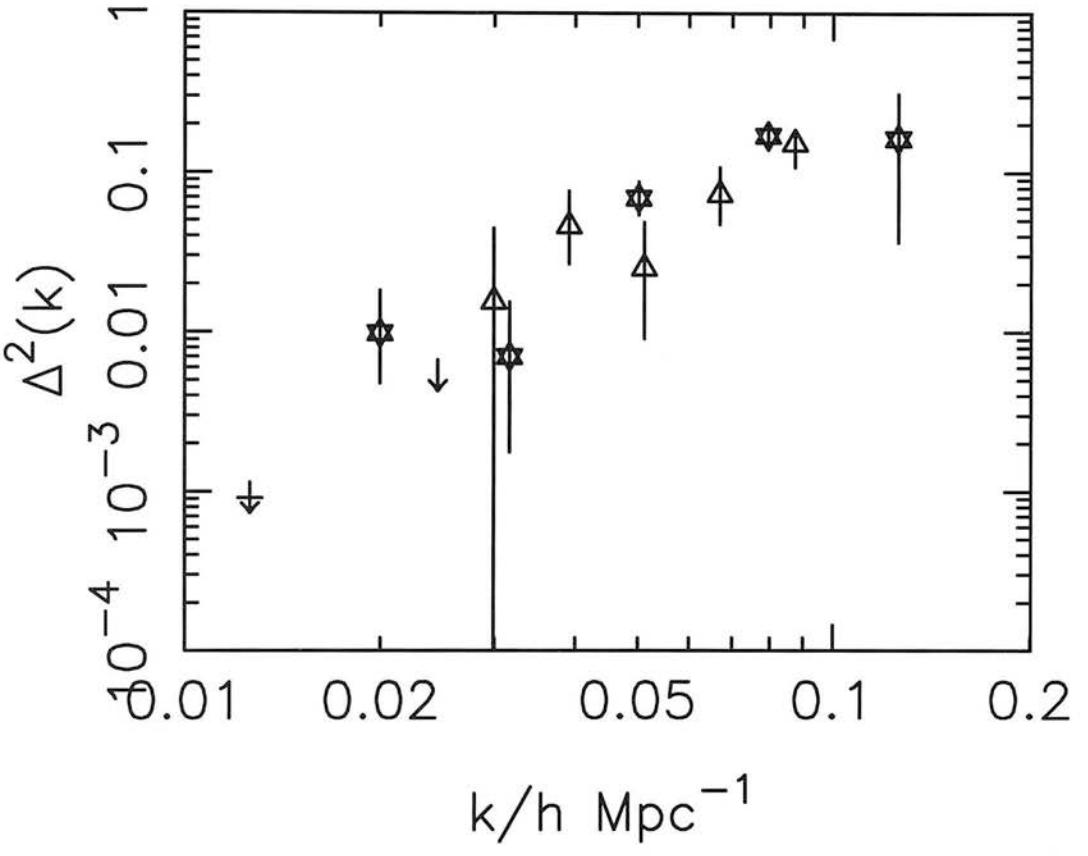

Figure 5.14: The real-space power spectrum of the PSCz (stars and cross-arrow) and 1.2Jy (triangles and arrow) redshift surveys.

haps possible that the final point is suppressed by nonlinearities, although the random smoothing model was used. A particular point to note is the very tight upper limit on the first bin, corresponding to rms fluctuations of only 3% on scales of $500\,h^{-1}\,\mathrm{Mpc}$, demonstrating the superb large-scale uniformity of the PSCz sample. The results still don't quite go deep enough to accurately constrain the expected primordial slope

Comparing the power spectra for the two surveys (Fig 5.14), a few things become clear. The PSCz reaches a greater depth, as expected, and constrains the power spectrum more accurately. There is also a slight difference in power normalisation of about 25% at $k \simeq 0.9\,h\,\mathrm{Mpc}^{-1}$, but this is about the difference expected with the new lower β value if the two surveys have the same *redshift-space* power spectra, see equation (2.18), so in that sense the results are consistent.

k/h Mpc^{-1}	$\Delta^2(k)$	$\Delta^2(k)$ lower	$\Delta^2(k)$ upper
0.0125	0.0	0.0	0.00092
0.0200	0.0099	0.0048	0.018
0.0316	0.0070	0.0018	0.015
0.0502	0.069	0.054	0.088
0.0795	0.17	0.15	0.20
0.126	0.16	0.036	0.31

Table 5.4: Power spectrum measurements and limits for the PSCz Survey.

5.4 Conclusions

For the first time, an estimation of the real-space power spectrum of *IRAS* galaxies on large scales $\lambda > 63h^{-1}$ Mpc has been determined, taking into account the spherical distortion introduced by peculiar velocities. We have performed simultaneous maximum likelihood fits for the real-space power spectrum and $\beta \equiv \Omega_0^{0.6}/b$ for the *IRAS* 1.2-Jy galaxy redshift survey. The relatively shallow depth of the survey precludes accurate measurement of $P(k)$ beyond 150 h^{-1} Mpc. We find $\beta = 1.04 \pm 0.3$, in agreement with the previous work of HT, who assumed a particular shape for the power spectrum.

The power spectrum for the 1.2Jy survey is noisy for scales $k < 0.03h$ Mpc^{-1}. It is consistent with a gradual turnover towards a Harrison-Zeldovich $P \propto k$ spectrum at around $k = 0.04$ h Mpc^{-1}, but a monotonically-rising power spectrum to small k cannot be ruled out with this method and dataset. It fits acceptably a CDM-type spectrum with shape parameter $\Gamma \simeq 0.2$, but *COBE*-normalized mixed dark matter models with a neutrino fraction of 13% or more do not fit the data well, requiring a substantial bias.

The spectrum for the PSCz survey has a slightly higher normalisation, although

the reasons for this are unclear. Again it appears to be well fit by a low Γ CDM model, although there is perhaps a hint of the sharper turnover of an MDM model or the Peacock fitting function.

More N-body modelling may be desirable, particularly for the stepwise case. Also, different bias schemes could be used to test if the method remains robust (cf. Hatton & Cole 1997). The ability to use only the largest scale modes means that nonlinear bias is expected to affect spherical harmonic methods less than others, but this should still be tested.

The value of β has fallen in the new analysis and is more in line with those of smaller scale methods such as Cole, Fisher & Weinberg (1995) who found $\beta \simeq 0.5$ using the power spectrum multipole method. It is unfortunate that the maximum likelihood value is so central, as it may be difficult to rule out either very high or very low β models with confidence, although the $r_{\rm max} = 270\,h^{-1}\,$Mpc model does appear to rule out $\beta = 1$. It is reassuring, however, that the β results from the stepwise power spectrum method match those from the two-parameter method so well. Recent results using POTENT-type methods (section 1.5.1) still tend to give higher values – Sigad *et al.* (1997) find $\beta_{IRAS} = 0.89 \pm 0.12$, although this uses the older 1.2Jy survey which also gave a high value with spherical harmonics. Note also that this inconsistent with other recent results *e.g.* Willick *et al.* (1997) found $\beta_{IRAS} = 0.49 \pm 0.07$ from a maximum-likelihood fitting of raw Tully-Fisher data, using exactly the same dataset as Sigad *et al.* (1997). Either the error bars are underestimated, or some of these methods are biased.

These power spectra are over the largest scales permitted by the survey, but they have a relatively short baseline so the constraint on the models is still relatively weak – even though these spectra probe the turnover, the most interesting part. This reliable analysis of the very large scale modes could be combined with other real space clustering analyses from smaller scales such as the cross correlation method of Saunders *et al.* (1992) to give a long baseline. However, it would be preferable to analyse more of the spectrum using just this technique. Future redshift surveys will contain many more linear modes and so this method will be applicable to a greater k-range. This

would increase the size of the covariance matrices and hence increase the already heavy burden on CPU time. In this case, it would probably be well worth looking at the data compression methods of the next chapter.

In addition to the relatively short baseline for $P(k)$, the error bars on β are relatively large compared with some other claims, *e.g.* Sigad *et al.* (1997) . This is probably because the cutoff in k was chosen so conservatively to preserve linearity. This implies that these results are particularly reliable, but it would still be desirable to find a tighter constraint on β. It is certainly possible to push the analysis into the nonlinear regime by extending the k range, although this would involve depending more heavily on the nonlinear modelling. Ideally the dispersion $\sigma_{\rm p}$ would be fitted as an extra parameter, although this is difficult in practice – the covariance matrices would need to be reconvolved from scratch at each point in parameter space. An alternative would be to crudely marginalise over the $\sigma_{\rm p}$ by choosing three possible values and creating three corresponding sets of covariance matrices. A Gaussian (parabola in log space) could then be fitted to perform a quick and crude integration at each point. If the nonlinearity is still mild this simple calculation should suffice, although N-body tests would certainly be required to confirm this. Again it should be pointed out that the power-spectrum nature of the transform allows the extent of the nonlinearity to be controlled. Of course, extending $k_{\rm max}$ will increase the number of modes, so again compression would be advisable.

The quoted errors on the power spectra are not entirely satisfactory. The error bars on the bins correspond to conditional errors which do not account for correlations between parameters. It would be preferable to explore parameter space more carefully to establish proper marginal errors, although again the computational cost would increase. The errors are nevertheless probably more justified than the "inverse" errors used in CFW95, where the measured value is taken to be the mean (see chapter 2).

The most important thing that these results need is an independent estimate of the linear bias. It would then be possible to constrain Ω_0 and the real-space *dark matter* power spectrum with confidence. A promising possibility is using the bispectrum, the covariance of three Fourier modes, see Matarrese, Verde & Heavens (1997).

Chapter 6

DATA COMPRESSION

The size of datasets available to cosmologists is increasing at an enormous rate. Giant redshift surveys containing $\sim 10^6$ galaxies will become available in the next few years, and new satellite missions will extract very detailed maps of microwave background fluctuations. The extraction of cosmological information from these surveys will become a very difficult task. The likelihood methods used throughout this work are the ideal way to extract cosmological parameters and their errors from observed data, but they require the repeated inversion of huge matrices and potentially require unfeasible computing power. A dataset containing n elements corresponds to an $n \times n$ covariance matrix. Even the power spectrum analysis of the PSCz survey in chapter 5 required over a week of CPU time on a powerful workstation. The detailed modelling possible with so many more data means that there will be more parameters to fit. Future CMB experiments will be able to constrain around 11 cosmological parameters (Jungman *et al.* 1996) so with processor time increasing exponentially with parameter number, the need for some kind of data compression becomes urgent.

This chapter is concerned with the development and testing of a method suggested by Tegmark, Taylor & Heavens (1997) (TTH), whereby the set of n is reduced in size via a linear transformation which minimises the loss of information about the desired parameters. TTH developed a general rigorous method for compression when the covariance matrix is a (possibly nonlinear) function of a single parameter, and

suggested an extension which could optimise for several parameters at once. They tested the proposed multi-parameter method by calculating the expected *conditional* errors on parameters in a CMB model. In this chapter the method was tested more fully using likelihood analysis of PSCz simulations. The two parameter model of chapter 5 was used and the TTH method was found to be inadequate to deal with correlated parameters – the conditional errors are constrained, but the marginal errors increased significantly. However, a more promising alternative is developed here which appears to handle parameter correlations very well, and should prove promising as a method to deal with the glut of data expected in the near future. The methods discussed here are also shown in Taylor, Heavens, Ballinger & Tegmark (1997).

The method of TTH is based on the Karhunen-Loève (KL) eigenvalue problem (Karhunen 1947) which has been used by a number of authors for CMB maps (Bond 1995;Bunn 1995; Bunn & Sugiyama 1995) and redshift surveys (Vogeley 1995; Vogeley & Szalay 1996).

The transforms (Fourier and Spherical) used in previous sections are themselves compression techniques, they are an alternative to directly comparing model redshift surveys/CMB maps with data. See TTH for a discussion of this "pre-compression" and for a more general introduction to and review of optimal data compression.

6.1 Maximum likelihood

Throughout the work in this thesis, the method of maximum likelihood is employed to fit models to observed data. This technique has some very nice features, and is in some sense an optimal way to estimate parameters, as described in detail in TTH. In this section that work is briefly reviewed.

6.1.1 Errors and Optimum Properties

In the following analysis a generic data discrete set will be considered, which is a set of n real numbers x_1, x_2, ..., x_n, arranged in a vector $\mathbf{x}$, with a vector $\boldsymbol{\mu}$ of mean values. These numbers could be *e.g.* real and imaginary parts of spherical transform coefficients of a galaxy redshift survey as in §4 or pixel values from a microwave background map. In the context of a model, the data may be regarded as a set of random variables. The m model parameters are similarly arranged in a vector

$$\boldsymbol{\Theta} = (\theta_1, \theta_2, ..., \theta_m). \tag{6.1}$$

The apparent values of parameters depend on the data, so these are also random variables; hence the true parameter values are denoted $\boldsymbol{\Theta}_0$, and $\boldsymbol{\Theta}$ are the estimates.

A good statistical estimator has a number of desirable properties; it is unbiased,

$$\langle \boldsymbol{\Theta} \rangle = \boldsymbol{\Theta}_0, \tag{6.2}$$

and minimises the error bars,

$$\Delta\theta_i \equiv \left(\left\langle \theta_i^2 \right\rangle - \langle \theta_i \rangle^2 \right)^{1/2}. \tag{6.3}$$

Hence the ideal estimator is known as the BUE θ_i, the "Best Unbiased Estimator".

A fundamental quantity in this analysis is the *Fisher Information Matrix*,

$$F_{ij} \equiv \left\langle \frac{\partial^2 \mathcal{L}}{\partial\theta_i \partial\theta_j} \right\rangle, \tag{6.4}$$

where $\mathcal{L} = -\log L$, where $L(\mathbf{x}, \Theta)$ is the likelihood function. The Maximum Likelihood estimator is the parameter vector $\Theta_{\rm ML}$ which maximises the likelihood function. The Fisher matrix is the inverse of the parameter covariance matrix if the errors are Gaussian (not the *data* covariance matrix which is to be inverted to evaluate the Likelihood function). In general, it gives a measure of the expected errors on parameters.

Several theorems can be proved about the Likelihood function, showing that it is the ideal way to extract parameters (see Kenney & Keeping 1951 or & Kendall & Stuart 1969):

1. All unbiased estimators have $\Delta\theta_i \geq 1/(F_{ii})^{1/2}$ (the Cramér-Rao inequality).

2. If a BUE Θ exists, it is the ML estimator or a function thereof.

3. In the limit of increasing data, the ML estimator is asymptotically a BUE and the parameter distribution is asymptotically Gaussian.

For a given dataset and model, the Cramér Rao inequality sets an absolute lower limit on the error bar of a parameter. This is exact if the parameter errors are Gaussian and the parameter is not correlated with any others. In the presence of correlations, the error rises to $\Delta\theta_i \geq (\mathbf{F}^{-1})_{ii}^{1/2}$.

6.1.2 Gaussian Data

If the data has a Gaussian distribution[1], as expected for modes of the linear-regime density field, the likelihood function is given by

$$L(\boldsymbol{\Theta}|\mathbf{x}, H_{\mathbf{x}}) = (2\pi)^{-n/2}|\mathbf{C}(\boldsymbol{\Theta})|^{-1/2}\exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^t\mathbf{C}(\boldsymbol{\Theta})^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]. \tag{6.5}$$

For $\mathcal{L} = -\log L$ this gives

$$2\mathcal{L} = \ln\det\mathbf{C} + (\mathbf{x}-\boldsymbol{\mu})\mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})^t, \tag{6.6}$$

dropping the additive constant $n\ln[2\pi]$. In general both the mean vector $\boldsymbol{\mu}$ and the covariance matrix

$$\mathbf{C} = \langle(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^t\rangle \tag{6.7}$$

depend on the parameters. This can be recast as a matrix formula by defining the data matrix

$$\mathbf{D} \equiv (\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^t \tag{6.8}$$

and using the matrix identity $\ln\det\mathbf{C} = \mathrm{Tr}\ln\mathbf{C}$, Recasting equation (6.6) as

$$2\mathcal{L} = \mathrm{Tr}\left[\ln\mathbf{C} + \mathbf{C}^{-1}\mathbf{D}\right]. \tag{6.9}$$

[1]This is not the limiting case of large n mentioned earlier, when the distribution of the *parameters* becomes Gaussian.

In what follows, the comma notation for derivatives will be used *e.g.*

$$\mathbf{C}_{,i} \equiv \frac{\partial}{\partial \theta_i}\mathbf{C}. \tag{6.10}$$

Note that $\mathbf{C}_{,i}$ is a matrix with elements defined as

$$(C_{,i})_{kl} = \frac{\partial C_{kl}}{\partial \theta_i}. \tag{6.11}$$

Because $\mathbf{C}$ is symmetric for all parameter values, clearly all the derivatives $\mathbf{C}_{,i}$, $\mathbf{C}_{,ij}$, *etc.*, will also be symmetric matrices. Taking the derivative of equation 6.9 and using the matrix identities $(\mathbf{C}^{-1})_{,i} = -\mathbf{C}^{-1}\mathbf{C}_{,i}\,\mathbf{C}^{-1}$ and $(\ln \mathbf{C})_{,i} = \mathbf{C}^{-1}\mathbf{C}_{,i}$, we find

$$2\mathcal{L}_{,i} = \mathrm{Tr}\left[\mathbf{C}^{-1}\mathbf{C}_{,i} - \mathbf{C}^{-1}\mathbf{C}_{,i}\,\mathbf{C}^{-1}\mathbf{D} + \mathbf{C}^{-1}\mathbf{D}_{,i}\right]. \tag{6.12}$$

At the true parameter values we have $\langle \mathbf{x} \rangle = \boldsymbol{\mu}$ and $\langle \mathbf{x}\mathbf{x}^t \rangle = \mathbf{C} + \boldsymbol{\mu}\boldsymbol{\mu}^t$, which gives

$$\begin{cases} \langle \mathbf{D} \rangle & = \mathbf{C}, \\ \langle \mathbf{D}_{,i} \rangle & = 0, \\ \langle \mathbf{D}_{,ij} \rangle & = \boldsymbol{\mu}_{,i}\,\boldsymbol{\mu}_{,j}^t + \boldsymbol{\mu}_{,j}\,\boldsymbol{\mu}_{,i}^t. \end{cases} \tag{6.13}$$

Using this and equation (6.12), we obtain $\langle \mathcal{L}_{,i} \rangle = 0$. This is reassuring, the ML-estimate is correct on average in the sense that the average slope of the likelihood function is zero at the point corresponding to the true parameter values. Differentiating equation (6.12), we obtain

$$\begin{aligned} 2\mathcal{L}_{,ij} = \mathrm{Tr}[\; & - \mathbf{C}^{-1}\mathbf{C}_{,i}\,\mathbf{C}^{-1}\mathbf{C}_{,j} + \mathbf{C}^{-1}\mathbf{C}_{,ij} \\ & + \mathbf{C}^{-1}(\mathbf{C}_{,i}\,\mathbf{C}^{-1}\mathbf{C}_{,j} + \mathbf{C}_{,j}\,\mathbf{C}^{-1}\mathbf{C}_{,i})\mathbf{C}^{-1}\mathbf{D} \\ & - \mathbf{C}^{-1}(\mathbf{C}_{,i}\,\mathbf{C}^{-1}\mathbf{D}_j + \mathbf{C}_{,j}\,\mathbf{C}^{-1}\mathbf{D}_i) \\ & - \mathbf{C}^{-1}\mathbf{C}_{,ij}\,\mathbf{C}^{-1}\mathbf{D} + \mathbf{C}^{-1}\mathbf{D}_{,ij}]. \end{aligned} \tag{6.14}$$

Taking the mean to find the Fisher matrix (equation (6.4)) and using equation (6.13) and the trace identity $\mathrm{Tr}[\mathbf{AB}] = \mathrm{Tr}[\mathbf{BA}]$, a much simpler expression drops out

$$F_{ij} = \langle \mathcal{L}_{,ij} \rangle = \frac{1}{2}\mathrm{Tr}[\mathbf{A}_i\mathbf{A}_j + \mathbf{C}^{-1}\mathbf{M}_{ij}], \tag{6.15}$$

where we have defined the matrices $\mathbf{A}_i \equiv \mathbf{C}^{-1}\mathbf{C}_{,i} = (\ln \mathbf{C})_{,i}$ and $\mathbf{M}_{ij} \equiv \langle \mathbf{D}_{,ij} \rangle = \boldsymbol{\mu}_{,i}\,\boldsymbol{\mu}_{,j}^t + \boldsymbol{\mu}_{,j}\,\boldsymbol{\mu}_{,i}^t$. This result is derived by TTH, Bunn (1995) and Vogeley & Szalay (1996).

6.2 Optimal Data Compression: Karhunen-Loève Methods

As shown in the previous section, Likelihood analysis has some useful statistical properties. It can, however, be computationally expensive – equation (6.6) involves the inversion of an $n \times n$ matrix, where n is the number of data values. The time for this inversion increases as n^3, so it can rapidly become large as datasets increase. Assuming that there are m parameters and that the m-dimensional parameter space is sampled as a grid of side q (cf the 15×15 sampling of the space of Δ and β in chapter 5), the timescale goes as

$$\tau \propto q^m \times n^3. \tag{6.16}$$

In addition to the problem of increasing n itself, large datasets lead to elaborate modelling so the number of parameters goes up as well. Jungman *et al.* (1996) present an 11 parameter CMB model, and the stepwise power spectrum estimation method of chapter 5 could be extended to many more separate bins in the analysis on a much larger survey. The time taken rises exponentially with the number of parameters, so compression is even more important.

There are several possible reasons why compression is possible. The data may be highly correlated, and so there may be significant redundancy among the data points; in particular, CMB experiments oversample the beam size to prevent aliasing (see TTH). Some linear combinations of modes may be far noisier than others, and have a negligible effect on parameter constraints. Also, only certain linear combinations of the data may be needed to estimate each of the parameters of interest. The latter possibility suggests that different compressed datasets may be ideal for different parameters – a theme which was explored by TTH, whose work is illustrated in this section.

6.2.1 The Optimisation Problem

This section follows the analysis of TTH; see also Bunn (1995).

Consider a general linear transform of the data

$$\mathbf{y} = \mathbf{B}\mathbf{x}, \tag{6.17}$$

the n'-dimensional vector $\mathbf{y}$ is the new compressed data set and $\mathbf{B}$ is an arbitrary $n' \times n$ matrix, with $n' \leq n$. Transforming the mean and variance:

$$\begin{cases} \langle \mathbf{y} \rangle & = \mathbf{B}\boldsymbol{\mu}, \\ \langle \mathbf{y}\mathbf{y}^t \rangle - \langle \mathbf{y} \rangle \langle \mathbf{y} \rangle^t & = \mathbf{B}\mathbf{C}\mathbf{B}^t, \end{cases} \tag{6.18}$$

and substituting this into equation (6.15) we find that the new Fisher information matrix $\tilde{\mathbf{F}}$ is given by

$$\begin{aligned} \tilde{F}_{ij} &= \frac{1}{2}\mathrm{Tr}[(\mathbf{B}\mathbf{C}\mathbf{B}^t)^{-1}(\mathbf{B}\mathbf{C}_{,i}\mathbf{B}^t)(\mathbf{B}\mathbf{C}\mathbf{B}^t)^{-1}(\mathbf{B}\mathbf{C}_{,j}\mathbf{B}^t) \\ &+ (\mathbf{B}\mathbf{C}\mathbf{B}^t)^{-1}(\mathbf{B}\mathbf{M}_{ij}\mathbf{B}^t)]. \end{aligned} \tag{6.19}$$

If $n = n'$ and $\mathbf{B}$ is invertible, then the $\mathbf{B}$-matrices cancel in this equation, leaving us with the original formula

$$\tilde{F}_{ij} = \frac{1}{2}\mathrm{Tr}[\mathbf{B}^{-t}(\mathbf{A}_i\mathbf{A}_j + \mathbf{C}^{-1}\mathbf{M}_{ij})\mathbf{B}^t] = F_{ij}. \tag{6.20}$$

This is not surprising, transforming by an invertible matrix – a similarity transformation – retains all the original information; it is always possible to transform back to the original matrices.

Compression can be achieved by having $n' < n$, *i.e.* y smaller than x. The new data vector is still a linear combination of elements of the old one, but information has been discarded. The crudest and simplest way to do this is simply to discard elements of x, making y a vector of a subset of the original data. This would seem to be far from ideal, and an apparent waste of a large dataset. However, it is often possible to achieve an almost lossless compression by asking what information is actually wanted from the data. This corresponds to picking a $\mathbf{B}$ matrix which minimises the error on a desired parameter for a given n'.

Considering a single row $\mathbf{b}$ of $\mathbf{B}$, the diagonal term of the Fisher matrix becomes

$$\tilde{F}_{ii} = \frac{1}{2}\left(\frac{\mathbf{b}^t\mathbf{C}_{,i}\,\mathbf{b}}{\mathbf{b}^t\mathbf{C}\mathbf{b}}\right)^2 + \frac{(\mathbf{b}^t\boldsymbol{\mu}_{,i})^2}{(\mathbf{b}^t\mathbf{C}\mathbf{b})}. \tag{6.21}$$

Concentrating on the problem of estimating a single parameter the off-diagonal correlation terms of the Fisher matrix can be ignored.

6.2.2 When the Mean is Known

In most of the cosmological areas where data is becoming excessive, *i.e.* the CMB and redshift surveys, the mean values of the data are known. For the more general case where this is not so see TTH (the additional analysis is very straightforward).

When $\boldsymbol{\mu}$ does not depend on θ_i, the second term in equation (6.21) vanishes, the quantity to be maximised becomes

$$(2\tilde{F}_{ii})^{1/2} = \frac{|\mathbf{b}^t \mathbf{C}_{,i}\, \mathbf{b}|}{\mathbf{b}^t \mathbf{C} \mathbf{b}}. \tag{6.22}$$

Since $\boldsymbol{\mu}$ is constant, it is convenient to redefine the data so that it has zero mean: $\mathbf{x} \rightarrow \mathbf{x} - \boldsymbol{\mu}$. Optimisation involves finding an extremum of $\mathbf{b}^t \mathbf{C}_{,i}\, \mathbf{b} / \mathbf{b}^t \mathbf{C} \mathbf{b}$ – not necessarily a maximum because the denominator of the rhs of equation (6.22) may be negative. Multiplying $\mathbf{b}$ by a constant does not change the ratio, so the denominator may be set to unity with no loss of generality. This allows a Lagrangian multiplier approach, extremising $\mathbf{b}^t \mathbf{C}_{,i}\, \mathbf{b}$ subject to the constraint that

$$\mathbf{b}^t \mathbf{C} \mathbf{b} = 1. \tag{6.23}$$

Introducing a Lagrange multiplier λ, an extremum of the following equation is sought,

$$\mathbf{b}^t \mathbf{C}_{,i}\, \mathbf{b} - \lambda \mathbf{b}^t \mathbf{C} \mathbf{b}. \tag{6.24}$$

The introduction of this constraint prevents the analysis finding the trivial solution that $\mathrm{d}\tilde{\mathbf{F}}/\mathrm{d}\mathbf{b} = 0$ if $\tilde{\mathbf{F}} = 0$. Differentiating with respect to the vector $\mathbf{b}$ (remembering $\mathbf{C}$ and $\mathbf{C}_{,i}$ are symmetric) gives the generalised eigenvalue problem for KL

$$\mathbf{C}_{,i}\, \mathbf{b} = \lambda \mathbf{C} \mathbf{b}. \tag{6.25}$$

This equation was actually derived by Bunn (1995), but was only used for *signal-to-noise eigenmodes*, where the covariance matrix is a linear function of a single parameter:

$$\mathbf{C} = \theta \mathbf{S} + \mathbf{N}, \tag{6.26}$$

where $\mathbf{S}$ is the signal matrix and $\mathbf{N}$ is the noise matrix – see also Bond (1995). Equation (6.25) then becomes simply

$$\mathbf{S} \mathbf{b} = \lambda \mathbf{C} \mathbf{b}. \tag{6.27}$$

This approach misses the power of equation (6.25), which can be used for *any* parameter, provided the derivative can be found.

The generalised eigenvalue problem can be solved by standard library routines. Since $\mathbf{C}$ is symmetric and positive definite, the very stable method of Cholesky decomposition can be used, see *e.g.*, Press *et al.* (1992). Although the standard routines can be regarded as black boxes, following the Cholesky formalism through demonstrates some nice properties of the transformed data. The covariance matrix is factored as $\mathbf{C} = \mathbf{L}\mathbf{L}^t$ for some invertible matrix $\mathbf{L}$ which can be regarded as the 'square root' of $\mathbf{C}$. Multiplying equation (6.25) by $\mathbf{L}^{-1}$ from the left converts it into an ordinary eigenvalue problem

$$(\mathbf{L}^{-1}\mathbf{C}_{,i}\,\mathbf{L}^{-t})(\mathbf{L}^t\mathbf{b}) = \lambda(\mathbf{L}^t\mathbf{b}), \tag{6.28}$$

The matrix $(\mathbf{L}^{-1}\mathbf{C}_{,i}\,\mathbf{L}^{-t})$ is symmetric, and there will be n orthogonal eigenvectors $(\mathbf{L}^t\mathbf{b}_k)$ with corresponding eigenvalues λ_k, $k = 1, 2, ..., n$. these can be sorted

$$|\lambda_1| \geq |\lambda_2| \geq ... \geq |\lambda_n|. \tag{6.29}$$

Let us choose the k^{th} row of $\mathbf{B}$ to be the row vector $\mathbf{b}_k^t$, so that the compressed data set is given by $y_k = \mathbf{b}_k^t\mathbf{x}$. The orthogonality property $(\mathbf{L}^t\mathbf{b}_k)\cdot(\mathbf{L}^t\mathbf{b}_{k'}) \propto \delta_{kk'}$ in combination with our chosen normalization of equation (6.23) then tells us that our compressed data satisfies

$$\langle y_k y_{k'}\rangle = \langle(\mathbf{b}_k^t\mathbf{x})(\mathbf{x}^t\mathbf{b}_{k'})\rangle = \mathbf{b}_k^t\mathbf{C}\mathbf{b}_{k'} = \mathbf{b}_k^t\mathbf{L}\mathbf{L}^t\mathbf{b}_{k'} = \delta_{kk'}, \tag{6.30}$$

i.e., $\langle\mathbf{y}\mathbf{y}^t\rangle = \mathbf{I}$ – new transformed covariance matrix is just the identity matrix. This would appear to make matrix inversion trivial, but this is only true at the point in parameter space where the Fisher matrix term is evaluated. Because the covariance matrix is now diagonal, the elements of $\mathbf{y}$ are independent[2]. The original data has been divided into n independent parts.

If all n vectors are retained, $\mathbf{B}$ will be square and invertible and no compression will have been done. Some of the $\mathbf{b}$ vectors will have to be discarded to reduce the size of $\mathbf{y}$.

[2]The new data are only independent at the assumed maximum likelihood point itself, so this does not confer a huge advantage, except that iterative inversion techniques may be faster near the fiducial point.

To decide the ideal way to choose which **b** vectors to discard, the expected error on the parameter can be expressed in terms of the eigenvalues λ. Combining all the vectors into a matrix **B** equation (6.25) becomes

$$\mathbf{C}_{,i}\,\mathbf{B}^t = \mathbf{C}\mathbf{B}^t\boldsymbol{\Lambda}, \tag{6.31}$$

where $\Lambda_{ij} \equiv \delta_{ij}\lambda_i$, *i.e.*, a diagonal matrix with the eigenvalues along the diagonal. Since the transformed covariance matrix is just the identity,

$$\mathbf{B}\mathbf{C}\mathbf{B}^t = \mathbf{I}, \tag{6.32}$$

multiplying equation (6.31) on the left by **B** gives

$$\mathbf{B}\mathbf{C}_{,i}\,\mathbf{B}^t = \boldsymbol{\Lambda}. \tag{6.33}$$

At these chosen parameter values, the diagonal element of equation (6.19) reduces to simply

$$2\tilde{F}_{ii} = \mathrm{Tr}[\{(\mathbf{B}\mathbf{C}\mathbf{B}^t)^{-1}(\mathbf{B}\mathbf{C}_{,i}\,\mathbf{B}^t)\}^2] = \mathrm{Tr}\boldsymbol{\Lambda}^2 = \sum_{k=1}^{n'} \lambda_k^2. \tag{6.34}$$

The compression is achieved by discarding **b** vectors corresponding to the smallest eigenvalues and working upwards. The eigenvalues should ordered so that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. If all the vectors are used ($n' = n$), **B** is invertible and there is no compression:

$$2\tilde{F}_{ii} = \sum_{k=1}^{n} \lambda_k^2 = 2F_{ii}. \tag{6.35}$$

Equation (6.34) shows that a single mode corresponding to $\mathbf{b}_i$ is expected to provide an error bar of $\Delta\theta_i = 1/|\lambda_i|$. TTH suggest that $\mathbf{b}_i$ vectors corresponding to modes of low "signal-to-noise", *i.e.* with $\theta_i|\lambda_i|$ below a certain threshold, should be discarded. However, it is possible to chose a value of n' in such a way that the effect on the parameter estimate is controlled more directly. The tolerance of the increase in the error bar can be fixed by choosing $n' < n$ such that

$$1 - \frac{\tilde{F}_{ii}}{F_{ii}} \leq \epsilon, \tag{6.36}$$

where, presumably, $\epsilon \ll 1$ in most cases.

The linear algebra leading to the KL transformations involves manipulating $n \times n$ matrices and, of course, this takes at least as much computing power as the calculation

of a likelihood using the full dataset. However, to explore parameter space requires *repeated* matrix inversions rather than just one, so the saving can be enormous (especially if the number of parameters is large – see equation (6.16)).

6.2.3 Estimating several parameters at once

In the previous section the tractable problem of optimising for one parameter was dealt with rigorously. However, most problems with large datasets involve estimating several parameters at once. If the likelihood problem is defined on a grid then the computation time rises exponentially with the number of parameters (see equation (6.16)) so data compression is particularly important in multi-parameter cases.

A simple solution suggested by TTH was to calculate the compression vectors separately for each parameter and then combine them together with an extra compression step. If the KL method is simply repeated m times, once for each parameter θ_i. If each θ_i has a corresponding n_i vectors then the resulting transformation matrix becomes very large, with $\mathbf{y}$ now a vector of length

$$n'' = \sum_{i=1}^{m} n_i'. \tag{6.37}$$

This analysis ignores the fact that the information in the data which estimates different parameters may overlap, so there could be a lot of redundancy in the new long $\mathbf{y}$. Indeed it may be longer than $\mathbf{x}$, a pointless anticompression producing a large, singular covariance matrix.

The problem can be thought of geometrically. Each set of vectors spans an n_i' dimensional subspace of the space where the data vector $\mathbf{x}$ exists. A lot of these subspaces overlap, so there basis vectors can be redundant. A new set of vectors is needed to economically span the combined subspace. This unnecessary overlap between vectors can be removed by making a *Singular Value Decomposition* (SVD) of the bloated $\mathbf{B}$, see Press *et al.* (1992). SVD constructs an orthonormal basis for the space spanned by the vectors in $\mathbf{B}$, with singular values giving the 'significance' of each vector *i.e.* very small singular values correspond to basis vectors which contribute very little to the estimation of any parameter. The vectors are weighted by their eigenvalues, but

unlike the TTH suggestion are rescaled so that each parameter gets the same total weight – a matrix $\mathbf{B}_i = \mathbf{B}\boldsymbol{\Lambda}'$ for each parameter. Without rescaling, parameters with low eigenvalues are further disadvantaged; clearly a parameter is no less important just because its error bar is large, if anything the opposite is true. Combining the new matrices into a single one and decomposing using SVD gives

$$\left(\mathbf{B}_1^t \ \ldots \ \mathbf{B}_m^t\right) = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^t, \tag{6.38}$$

where $\mathbf{U}^t\mathbf{U} = \mathbf{I}$, $\mathbf{V}^t\mathbf{V} = \mathbf{I}$ and the matrix $\boldsymbol{\Lambda}$ is diagonal. The diagonal elements of $\boldsymbol{\Lambda}$ are the singular values, and the columns of $\mathbf{U}$ contain the new basis vectors. Vectors corresponding to small singular values are discarded, in a similar way to those corresponding to small eigenvalues in the original KL analysis.

The final compression matrix for the TTH method is

$$\mathbf{B} \equiv \mathbf{U}^t \tag{6.39}$$

giving a new compressed data vector $\mathbf{z} = \mathbf{B}\mathbf{x}$. The columns of $\mathbf{U}$ span almost the entire subspace spanned by the combined $\mathbf{B}$ matrix – the entire subspace in the extreme of only discarding vectors corresponding to singular values which are exactly zero. The compressed data can again be made statistically orthogonal by Cholesky decomposing the new covariance matrix $\mathbf{BCB}^t = \mathbf{LL}^t$ and transforming again to $\mathbf{z}' = \mathbf{L}^{-t}\mathbf{z}$.

$$\langle \mathbf{y}\mathbf{y}^t \rangle = (\mathbf{U}^t\mathbf{L}^{-1})\mathbf{L}\mathbf{L}^t(\mathbf{U}^t\mathbf{L}^{-1})^t = \mathbf{I}, \tag{6.40}$$

i.e., that the compressed data set will be statistically orthogonal as before (although only at one point in parameter space).

6.3 Correlated Parameters

In many parameter estimation problems the parameters are correlated, $\langle(\theta_i - \mu_i)(\theta_j - \mu_j)\rangle \neq 0$ in general for $i \neq j$. This can lead to problems with the above analysis.

6.3.1 The Problem

The SVD method allows the optimisation of several parameters at once, but it is still finding the maximum value of diagonal Fisher matrix elements F_{ii}. This is equivalent to constraining the *conditional* errors on the parameters, *i.e.* the errors on each parameter if all the other parameters are held fixed. In practical cases, parameters are often highly correlated, and the *marginal* errors are dominated by this correlation. If the parameter distribution can be approximated by a multivariate Gaussian, the marginal error on a parameter is given by

$$\Delta\theta_i = (\mathbf{F}^{-1})_{ii}^{1/2} \tag{6.41}$$

instead of the previous formula $\Delta\theta_i = F_{ii}^{-1/2}$ which actually gives the conditional error; see Press *et al.* (1992, §15).

It may be possible to increase parameter independence by combining different datasets (*e.g.* redshift surveys and CMB maps), but inevitably, some correlations will remain.

The beauty of equation (6.25) is that it gives an algebraic solution which can be found reasonably quickly. There appears to be no such solution known for minimising $(\mathbf{F}^{-1})_{ii}^{1/2}$, so a complete analytic solution is not possible. It may be possible to eventually find the optimum $\mathbf{B}$ matrix using a numerical method, but this would probably take longer than the direct parameter estimation with the full covariance matrix. In the next section a useful practical technique is presented.

6.3.2 Parameter Eigenvector Optimisation

If the correlation is very large, marginal errors are dominated by the length of the ellipse, rather than by the conditional errors. This is clearly illustrated in Fig 6.1, the size of the marginal error MM' clearly depends more on the length of the ellipse rather than on the conditional interval CC'. In this case, constraining the F_{ii} elements directly seems somewhat pointless as the ellipse can grow along the primary principal axis (the "correlation axis") without changing the conditional errors, but greatly increasing the

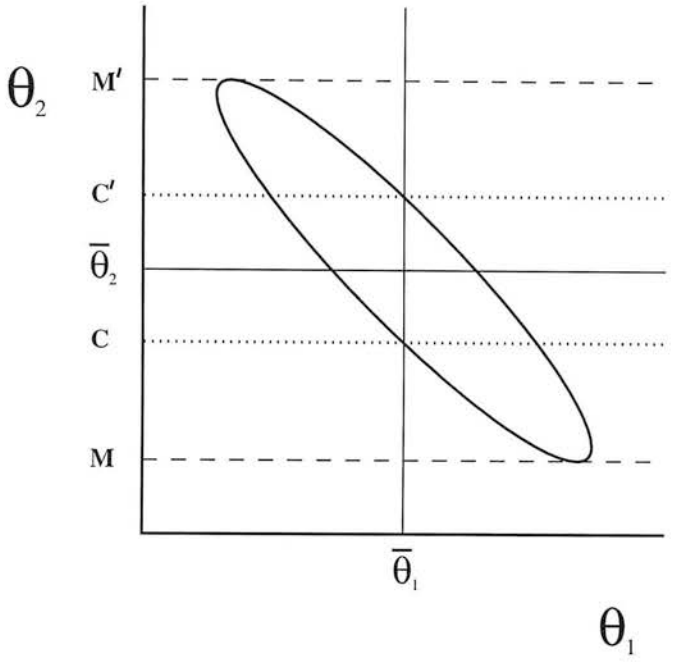

Figure 6.1: An artificial (mean) likelihood contour for two highly correlated parameters. The conditional error on parameter θ_2 is given by the interval CC', the marginal error by MM'; see Press *et al.* (1992). The mean values are denoted by $\bar{\theta_1}$ and $\bar{\theta_2}$.

variances. A better strategy is to constrain the error ellipse directly along this principal axis. This direction in parameter space corresponds to the eigenvector of the Fisher matrix with the smallest eigenvalue. Diagonalising,

$$\mathbf{F} = \mathbf{V}\mathbf{T}\mathbf{V}^t, \tag{6.42}$$

with

$$\mathbf{T} = \mathrm{diag}[t_1, t_2, \ldots], \tag{6.43}$$

where t_i is the i^{th} eigenvalue of $\mathbf{F}$. The rows of the $\mathbf{V}$ matrix correspond to the principal axes of the ellipse (or hyperellipsoid if there are more parameters) – see Arfken (1985). The diagonalisation of the Fisher matrix can be regarded as a rotation to a new parameter space, with n uncorrelated parameters which are linear combinations of the original ones[3],

$$\theta_i' = V_{ji}\theta_j. \tag{6.44}$$

[3]The new parameters Θ' are of course only used to find a useful compression matrix $\mathbf{B}$, the object of the analysis is still to constrain the original model parameters Θ.

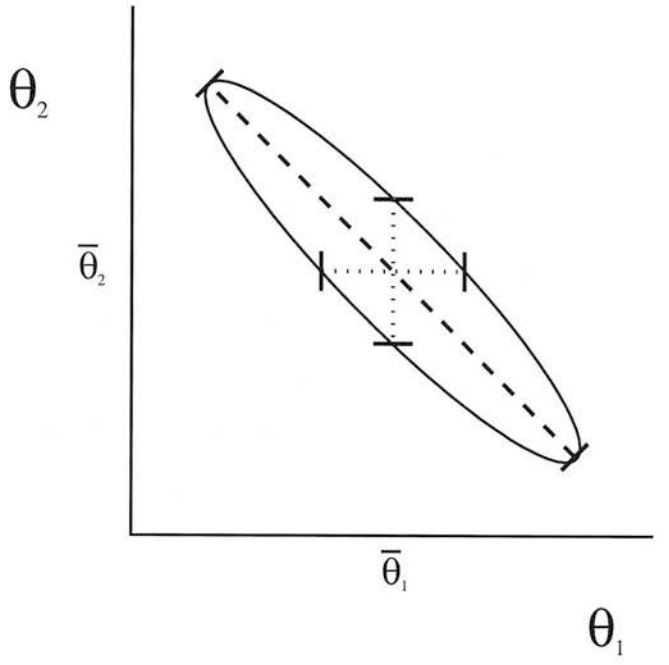

Figure 6.2: Illustration of the constraints imposed by different compression methods on the toy model of Fig 6.1. The dotted lines correspond to the conditional-error constraints imposed by the SVD method of TTH – clearly these have little influence on the length of the ellipse. The dashed line indicates the constraint imposed by the parameter-eigenvector method, a direct control of the length of the long axis itself.

The Fisher Matrix for these new parameters is the diagonal matrix $\mathbf{T}$ so, assuming Gaussian errors, the variances are given by

$$\mathbf{T}^{-1} = \mathrm{diag}[\sigma_1^2, \sigma_2^2, \ldots], \tag{6.45}$$

i.e. $\sigma_i = 1/\sqrt{t_i}$. Clearly there are no correlations between parameters and hence the TTH optimisation technique can be applied far more reliably. As can be seen from equation (6.45) the new parameter with the smallest Fisher eigenvalue has the largest error, and corresponds to the longest principal axis – the direction of the strongest correlation – and the second-smallest eigenvalue represents the next-longest axis *etc.* The derivatives are transformed in the same way as the vectors, as they are just gradients

$$\frac{\partial \mathbf{C}}{\partial \theta_i'} = V_{ji} \frac{\partial \mathbf{C}}{\partial \theta_j}, \tag{6.46}$$

so equation (6.31) can still be used without difficulty.

It may be worthwhile in practice to put some constraint on the shorter axes as well. In particular, if there are several axes of similar length, they should probably all be constrained. In this case, the transform matrices $\mathbf{B}$ for the different eigenvectors could be combined together using the SVD method of TTH to produce a simultaneous constraint. It is not clear if some general optimal way can be found to weight the vectors in the SVD depending on *e.g.* the relative strengths of the correlations – this awaits further work. The 11 parameter CMB model of Jungman *et al.* (1996b) will require a very extensive analysis keeping track of all the degeneracies between parameters (which will be stronger if there is significant reionisation – see TTH). The stages of a complete compression will probably involve:

1. A transform of the data – *e.g.* perform a spherical harmonic transform of the CMB, then average over coefficients with the same value of m, thus reducing the number of modes. See TTH for a discussion; they refer to this stage as "pre-compression".

2. Diagonalise the Fisher matrix to find the parameter eigenvectors.

3. Use equation (6.25) to find a $\mathbf{B}$ matrix for some or all of the transformed parameters – at least for those with small eigenvalues.

4. If more that one parameter eigenvector has been used, the SVD method of TTH should be used to combine the transform matrices together.

5. Trim the new $\mathbf{B}$ matrix, using the eigenvalues or singular values as described in section (6.2.2), fixing the accuracy using equation (6.36).

6. Transform the covariance matrix using the trimmed $\mathbf{B}$ matrix and carry out the likelihood analysis.

Note that the exact correspondence between the Fisher inverse and the covariances is strictly valid only if the parameters have a Gaussian distribution. The Fisher matrix really describes the local curvature of the likelihood surface (or hypersurface, in the case of more than 2 parameters).

6.4 Tests on Simulations

Fig (6.3) shows the effect of compression for the three techniques mentioned; optimising for a single parameter (β), the SVD method of TTH for both parameters and the new parameter eigenvector method. The full unrestricted set of modes is included for comparison. The data is from a simulation of the PSCz redshift catalogue, (see Saunders *et al.* 1995), analysed with the two-parameter spherical harmonic analysis of chapters 4 & 5. The range of uncompressed modes is restricted, with $l_{\max} = 16$, $n_{\max} = 6$, and $k_{\max} = 0.077\, h\,\mathrm{Mpc}^{-1}$, so the error bar is relatively large compared with what is attainable.

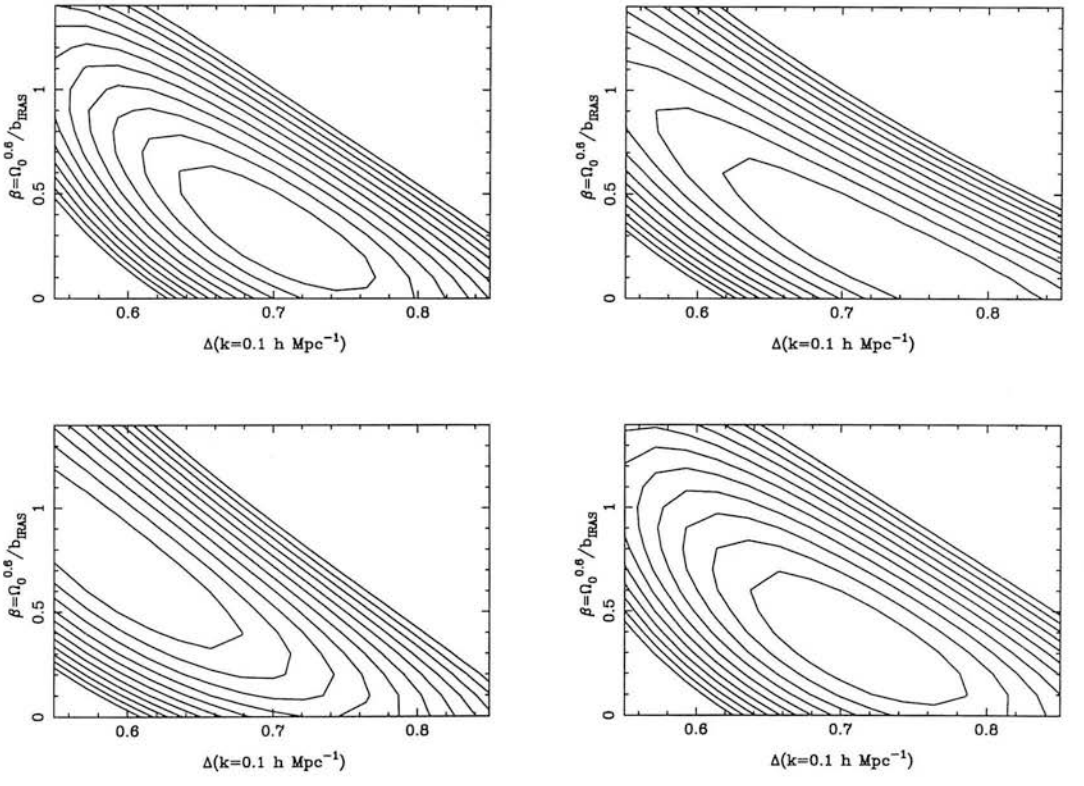

Figure 6.3: Illustration of data compression with different algorithms. Top left: 'Full' dataset of 508 modes (for details of parameters etc, see text). Top right: Best 320 modes optimised for measuring β. Bottom left: Best 320 modes from SVD application to modes optimised for β and Δ. Bottom right: Best 320 modes for optimising along the likelihood ridge axis. Likelihood contours are separated by 0.5 in natural log.

It is clear that the original TTH method is unsuitable for strongly correlated parameters. The conditional errors are still small, but the error ellipse is free to grow and shift along the correlation axis. The minimisation technique presumably throws away information about how to lift the degeneracy between the parameters. The eigenvector method explicitly prevents this behaviour by constraining the ellipse along that very axis.

It is a worry that even if the correlation is not large with the full dataset, it may increase considerably when only the conditional errors are constrained. The parameter eigenvector method should not be significantly more difficult to apply than its predecessors, so it would seem wise to use it in future compression analyses.

6.5 Discussion

Although this testing is preliminary, the parameter eigenvector technique seems to be ideal for the next generation of cosmological surveys. Hopefully, it will be developed and tested for future surveys and tame the gigantic datasets. Increasing numbers of parameters are going to be constrained by more and more data, creating impossible computer processor demands unless some compression can be performed. These models rarely seem to have clear, individual signatures for each parameter; instead they interact in a complicated fashion and strong correlations are almost inevitable.

It must be remembered that these methods are very general. Many problems both inside and outside astronomy make use of likelihood methods to fit parameters. Hopefully this technique will prove itself to be extremely useful in a wide range of applications, and allow otherwise intractable problems to be solved in a reasonable amount of time.

Chapter 7

CONCLUSIONS

This thesis is concerned with the extraction of cosmological information from galaxy redshift surveys; in particular, methods were developed for measuring the density parameter Ω_0, the cosmological constant Λ and the power spectrum $P(k)$. The former two parameters determine the fate of the universe, whether it will expand forever or recollapse; the power spectrum can constrain dark matter and inflationary models. The extraction of these parameters and statistics can be either helped or hindered by the redshift distortion effect – statistical changes in the redshift-space density field caused by peculiar velocities. A number of promising methods were developed and applied, with excellent results in some cases, and in others there are interesting prospects for analysis of future datasets:

1. An evolution-independent method for measuring Λ from forthcoming redshift surveys, accounting for the effects of linear and nonlinear redshift distortions. A full error analysis was performed.

2. A model-independent measurement of the large-scale *real-space* power spectrum was obtained from the 1.2Jy survey and the new PSCz survey, along with a simultaneous measurement of β from redshift distortions. A spherical transform was used to correctly deal with the geometry of the survey and the radial nature of the distortions.

3. A very general data-compression method was developed which can maintain the size of error bars on estimates of multiple *correlated* parameters while speeding up a statistical analysis. The method was tested and appears to be reliable. Such compression will be vital for analysing the giant datasets expected from new microwave-background and redshift surveys.

The analysis of redshift distortions is considered in detail in chapter 2. It has been shown that redshift distortions could have a significant effect on the observed density field, and cannot be ignored as it seems dangerous to assume that $\beta \ll 1$. The Kaiser analysis is very simple but may be unsuitable for present surveys due to its constraining assumptions of linearity and the distant observer approximation. The detailed linear modelling of Zaroubi & Hoffmann (1996) gets around the distant observer problem, but involves a huge 6D transformation and may introduce prohibitive computational costs, far clumsier than the 4D spherical harmonic analysis of chapters 4 and 5.

A simple nonlinear correction model for nonlinear effects was used to try to correct for the 'Fingers of God' effect in the context of the analysis of Cole, Fisher & Weinberg (1994). It was hoped that this would allow the method to be pushed to smaller scales. A simple incoherent distribution is added in to the Kaiser analysis, although there are doubts about the exact physical motivation for the model and the precise meaning of the parameter $\sigma_{\rm p}$, hence it seems best to fit this as a free parameter if possible. Quasi-linear infall has a similar effect to the random model so should also be modelled by it in practice. Random redshift errors correspond exactly to the 'Finger of God' model and the extra dispersion adds in quadrature. The model was found to be moderately successful at nonlinear modelling, as was also found by Cole, Fisher & Weinberg (1995). It should be quite robust and not depend on the detailed form of the distribution if the nonlinearity is mild – in fact it was shown that all distributions were the same to first order in $k^2\sigma_{\rm p}^2$. Current nonlinear models do not work in detail when the nonlinearity is significant, as was shown by Hatton & Cole (1997). It still seems desirable to go to larger scales where linearity is assured, or nonlinear effects are mild and can be modelled very simply.

As discussed in chapters 2 and 6, likelihood analysis is the ideal method for extracting parameters and finding realistic error bars including, for example, the β-dependence of errors. It is the statistical method employed throughout this thesis.

In chapter 3, an extra parameter was considered, the cosmological constant Λ. A number of authors had considered the measurement of this parameter (or, more generally, the deceleration parameter q_0) from measured structure at high redshift. This method takes advantage of the fact that the use of an incorrect cosmological model can lead to a geometric distortion – apparent anisotropic structure at high redshift. This can be used to give a robust method of measuring Λ depending purely on the assumption of isotropy but independent of the population and astrophysics of the high redshift objects used. For the first time, a full statistical analysis was performed which included all other effects, allowing a realistic analysis of the prospects of measuring Λ with future surveys. It was shown that there is a serious risk of confusion with redshift distortion anisotropies, and difficulties with the unknown evolution of bias; very large datasets are required to distinguish the two effects.

A detailed statistical model was developed which included linear and nonlinear redshift distortions and shot noise as well as the Λ geometric effect. The correlations between the parameters was investigated in detail, and in particular it was found that $\sigma_{\rm p}$ and the geometric effect are almost independent. The prospects for constraining Λ with the proposed 2dF galaxy and quasar surveys were investigated, and plausible error estimates were derived. The feasibility was found to be uncertain, and depends strongly on the actual values of other parameters and the exact realisation of the stochastic density field. Ironically, strong linear redshift distortions – high β – make the geometric signal easier to distinguish. This leads to a tradeoff, however, as lower bias leads to higher β but low power-spectrum amplitude *i.e.* poorer signal-to-noise. The results may also depend heavily on the nonlinear model and the simple "Finger of God" scatter may not be good enough; the signs are promising, but clearly detailed N-body modelling is needed.

In order to measure large-scale radial distortions in all-sky surveys (in particular the IRAS surveys) the spherical harmonic formalism of HT95 was employed and ex-

tended. A non-parametric measurement of the large scale *real-space* power spectrum was developed using the stepwise maximum likelihood technique, allowing simultaneous measurement of the power spectrum shape and β. In addition, a new all-sky IRAS redshift survey became available which was ideally suited to these spherical harmonic techniques.

The first large-scale real-space power spectrum of the IRAS 1.2Jy was found, and showed strong evidence of a turnover. The shape was well fitted by the two-power-law model of Peacock & Dodds (1994) (which was assumed in HT), and the value of $\beta \simeq 1$ consistent with the simple 2-parameter result of HT. The value of $\beta \sim 1$ was consistent with that found by Heavens & Taylor (1995)

The new PSCz survey is currently the largest near-all-sky survey that exists. An enhanced version of the HT analysis yielded a lower value of $\beta \simeq 0.6$, more in line with other results, although the POTENT values remain quite high: $\beta = 0.89 \pm 0.12$. This will require a biasing of IRAS galaxies to give $\Omega = 1$, although it fits well with currently popular open models if the galaxies are unbiased. The power spectrum results again show a turnover, with a remarkably strong lower limit in the final bin, corresponding to a large-scale smoothness in the catalogue of only a few percent. The power spectrum was well modelled by CDM, but is perhaps better fit by a model with a sharper break, such as MDM or the simple fit of Peacock & Dodds (1994). It is clear that for the PSCz, the analysis should be pushed further into the nonlinear regime, with "finger-of-God" effects dealt with more carefully. This is for two reasons, to drive down the error on β and to increase the baseline of the measured power spectrum to improve the model fitting. This will also increase the size of the covariance matrices, and probably warrants the use of the data-compression method of chapter 6.

The spherical harmonic approach to power-spectrum fitting seems ideal for extracting the power at the largest scales from future very large surveys. With this method, future surveys such as the Sloan survey should sample the primordial power spectrum slope, and this technique seems ideal for finding the undistorted shape.

In general, future datasets will become very large and Likelihood analysis of correlated data will involve enormous matrix calculations. It is vital that a data compression technique can be used to extract cosmological information. The parameter eigenvector optimisation method outlined in chapter 6 seems to provide an ideal practical method, easily extendable to future multiparameter problems. It is the first practical method which can deal with several correlated parameters at once, a vital feature in the analysis of future surveys. In particular, the proposed Planck Surveyor will produce so many data that any but the crudest analysis will require some sort of compression. The technique is very general and could have many applications inside and outside of astronomy.

In the past few years, with improving data and methods to analyse it, cosmology has become very exciting. Standard CDM ($\Omega_0 = 1$) appears to be in serious difficulties, and there are problems reconciling the ages of globular clusters with the age of an Einstein-de Sitter universe. Evolved galaxies are being seen at high redshift (Dunlop *et al.* 1996), putting further pressure on the $\Omega_0 = 1$ models which became a standard assumption after inflation was proposed. Open and Λ models are becoming popular, and even more exotic models have been suggested. Detection of high redshift galaxies is increasing rapidly, and a clearer picture of the evolution of galaxies is beginning to emerge (*e.g.* Madau *et al.* 1996). The evolution of structure is beginning to be analysed (see Peacock 1997), and this work will increase greatly in the future. In large scale structure power spectra have been measured and redshift distortions detected but inconsistencies remain, although tests for non-Gaussian effects do tend to be negative (see below). There is clearly a lot of work still to be done in this field.

This thesis has been concerned with the analysis of large scale structure, and there are perhaps two main uncertainties plaguing this field, Gaussianity and bias. The former is relatively benign, recent tests of this assumption have appear to support it, *e.g.* Feldmann, Kaiser & Peacock (1994) and Stirling & Peacock (1996), and popular theories of the early universe tend to predict it. Nevertheless, more subtle methods may find that it breaks down in detail. The problem of bias is the greatest bugbear of this subject, it is essentially a "fudge factor" to parameterise our ignorance of how galaxies trace mass. All linear redshift distortion analysis suffers from the degeneracy between

Ω and the bias parameter b. This could be lifted by finding b independently, using a nonlinear analysis such as the bispectrum analysis of Matarrese, Verde & Heavens (1997). It may also be possible to use the details of nonlinear redshift distortions to find b, as suggested by Cole *et al.* (1994). Linear biasing is usually assumed for simplicity, although there are now better reasons for believing it, at least at large scales. There is evidence that linear biasing breaks down on intermediate scales – see Peacock (1997); and it may be that the model is more complicated with stochastic biasing where bias itself is a random variable, adding extra structure. This is an important problem to be addressed by future surveys.

Cosmology is on the verge of a data explosion. New methods of taking redshifts are going to create "supersurveys", which will largely eclipse previous efforts. The work in this thesis straddles the two generations, with some methods applied to the current state-of-the-art, and others waiting for the next generation to come into there own. The PSCz is one of the last of the old generation of samples, taken from a variety of sources and telescopes without the advantage of new, fast fibre spectrographs. For some time to come it will probably remain the definitive guide to local cosmography, and the ideal survey for analysing the origin of our own motion through the universe. It was an ideal survey for spherical harmonic techniques, and, particularly with further refinement (an increase in $k_{\rm max}$) these techniques provide a good estimate of β_{IRAS}, competitive with other current estimates (and probably more reliable). There is no doubt, however, that the forthcoming supersurveys will measure β with far more precision (see chapter 3), and efforts will largely be spent on pursuing subtler effects such as evolution, biasing and, of course, the geometric signature of Λ. Nevertheless, the methods of chapters 4 and 5 will remain a competitive way to measure large-scale power. The measurement of Λ is definitely a task for the future, even the 2-degree Field surveys will struggle to constrain it. Although other methods such as supernovae are used to find q_0 (Perlmutter *et al.* 1997), they lack the purity of the large scale structure method with its independence of source evolution, and it warrants further study. Finally the data compression method devised in this thesis seems to be very timely, with obvious applications to the new giant redshift and CMB surveys and the multi-parameter models which will be tested. In the forthcoming deluge of data, the powerful methods contained in this

thesis will be extremely useful for finally solving many cosmological puzzles which have been pondered for decades.

Appendix A

Moments of the Power Spectrum Including Nonlinear Distortions

The power spectrum moments can be calculated analytically by direct integration of equation (2.24)[1]:

$$P_l^S(k) = \frac{2l+1}{2} \int_{-1}^{+1} \mathrm{d}\mu \, P^S(k,\mu) \mathcal{P}_l(\mu)$$

A.1 Moments for Gaussian Model

For the Gaussian model:

$$P^S(k,\mu) = P^R(k)(1+\beta\mu^2)^2 \exp(-\mu^2 A^2) \tag{A.1}$$

where:

$$A = \frac{1}{\sqrt{2}} \left(\frac{k\sigma_\mathrm{p}}{H} \right) \tag{A.2}$$

The $l = 0$ moment is given by:

$$P_0^S = \frac{1}{2} P^R(k) \int_{-1}^{+1} \mathrm{d}\mu \, (1+\beta\mu^2)^2 \exp(-\mu^2 A^2) \tag{A.3}$$

[1]For reference, the first two even Legendre polynomials are $\mathcal{P}_0 = 1$ and $\mathcal{P}_2 = (3\mu^2 - 1)/2$.

Expanding and integrating terms separately:

$$\begin{aligned}
\frac{1}{2}\int_{-1}^{+1} \mathrm{d}\mu\ \exp(-\mu^2 A^2) &= \frac{\sqrt{\pi}}{2A}\,\mathrm{erf}(A) \\
\beta\int_{-1}^{+1} \mathrm{d}\mu\,\mu^2\ \exp(-\mu^2 A^2) &= -\frac{\beta}{A^2}\exp(-A^2) + \frac{\sqrt{\pi}\beta}{2A^3}\,\mathrm{erf}(A) \\
\frac{\beta^2}{2}\int_{-1}^{+1} \mathrm{d}\mu\,\mu^4\ \exp(-\mu^2 A^2) &= -\frac{\beta^2}{2A^2}\exp(-A^2) - \frac{3\beta^2}{4A^4}\exp(-A^2) + \frac{3\sqrt{\pi}\beta^2}{8A^5}\,\mathrm{erf}(A)
\end{aligned} \tag{A.4}$$

Hence the $l = 0$ moment is:

$$\begin{aligned}
P_0^S(k,\mu) &= P^R(k)\left[-\left(\frac{(2\beta+\beta^2)}{2A^2} + \frac{3\beta^2}{4A^4}\right)\exp(-A^2)\right. \\
&+ \left.\left(\frac{\sqrt{\pi}}{2A} + \frac{\sqrt{\pi}\beta}{2A^3} + \frac{3\sqrt{\pi}\beta^2}{8A^5}\right)\mathrm{erf}(A)\right]
\end{aligned} \tag{A.5}$$

The $l = 2$ moment is given by:

$$P_0^S = \frac{5}{4}P^R(k)\int_{-1}^{+1} \mathrm{d}\mu\,(1+\beta\mu^2)^2(3\mu^2-1)\,\exp(-\mu^2A^2) \tag{A.6}$$

Expanding and integrating terms separately (not including $\frac{5}{4}$ factor):

$$\begin{aligned}
-\int_{-1}^{+1} \mathrm{d}\mu\ \exp(-\mu^2A^2) &= -\frac{\sqrt{\pi}}{A}\,\mathrm{erf}(A) \\
(3-2\beta)\int_{-1}^{+1} \mathrm{d}\mu\,\mu^2\ \exp(-\mu^2A^2) &= -\frac{(3-2\beta)}{A^2}\exp(-A^2) + \frac{\sqrt{\pi}(3-2\beta)}{2A^3}\,\mathrm{erf}(A) \\
(6\beta-\beta^2)\int_{-1}^{+1} \mathrm{d}\mu\,\mu^4\ \exp(-\mu^2A^2) &= -\frac{(6\beta-\beta^2)}{A^2}\exp(-A^2) - \frac{3(6\beta-\beta^2)}{2A^2}\exp(-A^2) \\
&+ \frac{3\sqrt{\pi}(6\beta-\beta^2)}{4A^5}\,\mathrm{erf}(A) \\
3\beta^2\int_{-1}^{+1} \mathrm{d}\mu\,\mu^6\ \exp(-\mu^2A^2) &= -\frac{3\beta^2}{A^2}\exp(-A^2) - \frac{15\beta^2}{2A^4}\exp(-A^2) \\
&- \frac{45\beta^2}{4A^6}\exp(-A^2) + \frac{45\sqrt{\pi}\beta^2}{8A^7}\,\mathrm{erf}(A)
\end{aligned} \tag{A.7}$$

Hence the $l = 2$ moment is:

$$\begin{aligned}
&P_2^S(k,\mu) = \\
&\frac{5}{4}P^R(k)\left[-\left(\frac{(3+4\beta+2\beta^2)}{A^2} + \frac{(9\beta+6\beta^2)}{A^4} + \frac{45\beta^2}{4A^6}\right)\exp(-A^2)\right. \\
&+ \left.\left(-\frac{\sqrt{\pi}}{A} + \frac{\sqrt{\pi}(3-2\beta)}{2A^3} + \frac{3\sqrt{\pi}(6\beta-\beta^2)}{4A^5} + \frac{45\sqrt{\pi}\beta^2}{8A^7}\right)\mathrm{erf}(A)\right]
\end{aligned} \tag{A.8}$$

A.2 Moments of exponential model

For the exponential model:

$$P^S(k,\mu) = P^R(k)\frac{(1+\beta\mu^2)^2}{(1+\mu^2A^2)} \tag{A.9}$$

where again:

$$A = \frac{1}{\sqrt{2}}\left(\frac{k\sigma_{\rm p}}{H}\right)$$

The $l=0$ moment is given by:

$$P_0^S = \frac{1}{2}P^R(k)\int_{-1}^{+1}\mathrm{d}\mu\,\frac{(1+\beta\mu^2)^2}{(1+\mu^2A^2)} \tag{A.10}$$

Expanding and integrating terms separately:

$$\begin{aligned}
\frac{1}{2}\int_{-1}^{+1}\mathrm{d}\mu\,\frac{1}{(1+\mu^2A^2)} &= \frac{1}{A}\tan^{-1}(A)\\
\int_{-1}^{+1}\mathrm{d}\mu\,\frac{\beta\mu^2}{(1+\mu^2A^2)} &= 2\beta\left(\frac{1}{A^2}-\frac{1}{A^3}\tan^{-1}(A)\right)\\
\frac{1}{2}\int_{-1}^{+1}\mathrm{d}\mu\,\frac{\beta^2\mu^4}{(1+\mu^2A^2)} &= \beta^2\left(\frac{1}{3A^5}-\frac{1}{A^2}+\frac{1}{A^5}\tan^{-1}(A)\right)
\end{aligned} \tag{A.11}$$

Hence the $l=0$ moment is:

$$P_0^S(k,\mu) = P^R\left[\frac{(\beta^2+6\beta)}{3A^2}-\frac{\beta^2}{A^4}+\left(\frac{1}{A}-\frac{2\beta}{A^3}+\frac{\beta^2}{A^5}\right)\tan^{-1}(A)\right] \tag{A.12}$$

The $l=2$ moment is given by:

$$P_0^S = \frac{5}{4}P^R(k)\int_{-1}^{+1}\mathrm{d}\mu\,\frac{(1+\beta\mu^2)^2(3\mu^2-1)}{(1+\mu^2A^2)} \tag{A.13}$$

Expanding and integrating terms separately (not including $\frac{5}{4}$ factor):

$$\begin{aligned}
-\int_{-1}^{+1}\mathrm{d}\mu\,\frac{1}{(1+\mu^2A^2)} &= -\frac{2}{A}\tan^{-1}(A)\\
\int_{-1}^{+1}\mathrm{d}\mu\,\frac{(3-2\beta)\mu^2}{(1+\mu^2A^2)} &= (6-4\beta)\left(\frac{1}{A^2}-\frac{1}{A^3}\tan^{-1}(A)\right)\\
\int_{-1}^{+1}\mathrm{d}\mu\,\frac{(6\beta-\beta^2)\mu^4}{(1+\mu^2A^2)} &= (12\beta-2\beta^2)\left(\frac{1}{3A^2}-\frac{1}{A^4}+\frac{1}{A^5}\tan^{-1}(A)\right)\\
\int_{-1}^{+1}\mathrm{d}\mu\,\frac{3\beta^2\mu^6}{(1+\mu^2A^2)} &= 6\beta^2\mu^6\left(\frac{1}{5A^2}-\frac{1}{3A^4}+\frac{1}{A^6}-\frac{1}{A^7}\tan^{-1}(A)\right)
\end{aligned} \tag{A.14}$$

Hence the $l = 2$ moment is:

$$
\begin{aligned}
&P_2^S(k,\mu) = \\
&\quad \frac{5}{4}P^R(k)\left[\frac{(90+8\beta^2)}{15A^2} - \frac{12\beta}{A^4} + \frac{6\beta^2}{A^6}\right. \\
&\quad \left. + \left(-\frac{2}{A} - \frac{(6-4\beta)}{A^3} + \frac{(12\beta-2\beta^2)}{A^5} - \frac{6\beta^2}{A^7}\right)\tan^{-1}(A)\right] \qquad \text{(A.15)}
\end{aligned}
$$

Appendix B

Fourier Transforms

The squashing effect transforms coordinates and wave vectors as follows:

$$\mathbf{x}' = \mathbf{S} \cdot \mathbf{x} \quad ; \quad \mathbf{k}' = \mathbf{S}^{-1} \cdot \mathbf{k}, \tag{B.1}$$

with transformation matrix

$$\mathbf{S} = \begin{pmatrix} f_\perp^{-1} & 0 & 0 \\ 0 & f_\perp^{-1} & 0 \\ 0 & 0 & f_\parallel^{-1} \end{pmatrix}. \tag{B.2}$$

The Λ squashing effect on the power spectrum is thus the following transform of the (possibly anisotropic) correlation function:

$$P(\mathbf{k}) = \int \xi(\mathbf{r})\, \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} \mathrm{d}^3 r, \tag{B.3}$$

$$P'(\mathbf{k}') = \int \xi(\mathbf{r})\, \mathrm{e}^{\mathrm{i}(\mathbf{S}\cdot\mathbf{k}')\cdot\mathbf{r}}\, |\mathbf{S}|\, \mathrm{d}^3 r = |\mathbf{S}|\, P(\mathbf{S}\cdot\mathbf{k}'). \tag{B.4}$$

Note that $\xi'(\mathbf{r}') = \xi(\mathbf{r})$ since $\delta'(\mathbf{r}') = \delta(\mathbf{r})$. Apart from an amplitude shift the original power spectrum is retained, but evaluated at $\mathbf{S}\cdot\mathbf{k}'$:

$$P'(k'_\parallel, \mathbf{k}'_\perp) = \frac{1}{f_\perp^2 f_\parallel} P\left(\frac{k'_\parallel}{f_\parallel}, \frac{\mathbf{k}'_\perp}{f_\perp}\right). \tag{B.5}$$

Redshift distortions are modelled as a product of a Kaiser factor and a damping term

$$\begin{aligned} P(k_\parallel, \mathbf{k}_\perp) &= P_0(k)\, (1+\beta\mu^2)^2\, D(k\mu\sigma_\mathrm{p}) \\ &= P_0(k)\, k^{-4}\, [k_\perp^2 + (\beta+1)k_\parallel^2]\, D(k_\parallel\sigma_\mathrm{p}), \end{aligned} \tag{B.6}$$

where $D(k\mu\sigma_{\rm p})$ is the non-linear 'finger of God' correction. Combining redshift and Λ effects gives

$$\begin{aligned} P'(\mathrm{k}') &= \frac{1}{f_\perp^2 f_\parallel} P_0 \left(\sqrt{\frac{k_\perp'^2}{f_\perp^2} + \frac{k_\parallel'^2}{f_\parallel^2}} \right) \left(\frac{k_\perp'^2}{f_\perp^2} + \frac{k_\parallel'^2}{f_\parallel^2} \right)^{-2} \\ &\times \left[\frac{k_\perp'^2}{f_\perp^2} + (\beta+1) \frac{k_\parallel'^2}{f_\parallel^2} \right] D\left(\frac{k_\parallel \sigma_{\rm p}}{f_\parallel} \right). \end{aligned} \tag{B.7}$$

Now we set $F = f_\parallel / f_\perp$, $k_\parallel' = \mu' k'$ and $\sigma_{\rm p}' = \sigma_{\rm p}/f_\parallel$, which yields

$$\begin{aligned} P'(\mathrm{k}') &= \frac{1}{f_\perp^2 f_\parallel} P_0 \left[\frac{k'}{f_\perp} \sqrt{1 + \mu'^2 \left(\frac{1}{F^2} - 1 \right)} \right] \\ &\times \left[1 + \mu'^2 \left(\frac{1}{F^2} - 1 \right) \right]^{-2} \\ &\times \left[1 + \mu'^2 \left(\frac{\beta+1}{F^2} - 1 \right) \right]^2 D\left(k'\mu'\sigma_{\rm p}' \right). \end{aligned} \tag{B.8}$$

For a power spectrum that is locally close to a power law, with index $n = \mathrm{d}\ln P / \mathrm{d}\ln k$, we have

$$\begin{aligned} P'(\mathrm{k}') &= \frac{1}{f_\perp^{3+n} F} P(k') \left[1 + \mu'^2 \left(\frac{1}{F^2} - 1 \right) \right]^{\frac{n-4}{2}} \\ &\times \left[1 + \mu'^2 \left(\frac{\beta+1}{F^2} - 1 \right) \right]^2 D\left(k'\mu'\sigma_{\rm p}' \right). \end{aligned} \tag{B.9}$$

Assuming an exponential distribution for the random pairwise velocity component, the non-linear correction to the power spectrum is a Lorentzian factor:

$$D\left(k'\mu'\sigma_{\rm p}' \right) = \frac{1}{1 + \left(k'\mu'\sigma_{\rm p}' \right)^2 / 2} \tag{B.10}$$

(where the units of $\sigma_{\rm p}'$ are h^{-1}Mpc).

Appendix C

Selection functions and Weighting schemes

The number of independent modes depends on the details of the survey. We derive here the effective density of states for a varying selection function and weighting of the data. The simplest case is a uniformly sampled cube of side L, containing $N_{\rm g}$ galaxies, where independent modes are separated by $\Delta k = 2\pi/L$ in k-space, so that

$$\rho(k) = \left(\frac{L}{2\pi}\right)^3. \tag{C.1}$$

In this case (e.g. Peacock & Nicholson 1991), one would define Fourier coefficients via

$$\delta_{\mathbf{k}} = \frac{1}{N_{\rm g}} \sum \exp[{\rm i}\mathbf{k}\cdot\mathbf{x}], \tag{C.2}$$

and in the continuum limit we would have

$$\langle L^3 \, |\delta_{\mathbf{k}}|^2 \rangle = P_{\rm true} + P_{\rm shot} \tag{C.3}$$

where the shot noise depends on the density $n = N_{\rm g}/L^3$:

$$P_{\rm shot} = \frac{1}{n}. \tag{C.4}$$

In these expressions, we use the Fourier convention of Peebles (1980), but with unit normalization volume:

$$P(k) = \int \xi(r) \, \exp[{\rm i}\mathbf{k}\cdot\mathbf{r}] \, {\rm d}^3 r. \tag{C.5}$$

These relations show that there is some advantage to having a large survey, thus giving a higher density of states. However, for a fixed number of galaxies, increasing the volume increases the shot noise on each mode. There is an optimum survey size for a given number of galaxies which corresponds to a signal-to-noise ratio of about unity.

The general case of a survey of non-uniform sampling was considered by Feldman, Kaiser & Peacock (1994; hereafter FKP). They also allowed the galaxy density field to be weighted so as to optimize the signal-to-noise ratio in the power spectrum:

$$\delta_{\mathbf{k}} = \frac{\int w[n - \bar{n}]\ \exp[\mathrm{i}\mathbf{k}\cdot\mathbf{r}]\ \mathrm{d}^3r}{[\int w^2\bar{n}^2\ \mathrm{d}^3r]^{1/2}}, \tag{C.6}$$

where $\bar{n}(\mathbf{r})$ is the mean density defined by the sample selection. With this generalization,

$$\langle|\delta_{\mathbf{k}}|^2\rangle = P_{\mathrm{true}} + P_{\mathrm{shot}} \tag{C.7}$$

where the shot noise is now

$$P_{\mathrm{shot}} = \frac{\int w^2\bar{n}\ \mathrm{d}^3r}{\int w^2\bar{n}^2\ \mathrm{d}^3r}. \tag{C.8}$$

FKP also derive (their equation 2.3.2) the ratio of the variance in the total power to P^2_{true}. For a uniform cubical survey of side L, this would be

$$\frac{\sigma(P)}{P_{\mathrm{tot}}} = N^{-1/2}_{\mathrm{modes}} = [\rho_k V_k]^{-1/2}, \tag{C.9}$$

which is proportional to $L^{-3/2}$. Converting the FKP result to the ratio of total variance to total power gives an effective survey size of

$$L^{-3}_{\mathrm{eff}} = \frac{\int w^4\bar{n}^4[P + \bar{n}^{-1}]^2\ \mathrm{d}^3r}{[P\int w^2\bar{n}^2\ \mathrm{d}^3r + \int w^2\bar{n}\ \mathrm{d}^3r]^2.} \tag{C.10}$$

This reduces to $L_{\mathrm{eff}} = L$ for the previous case of a uniform cube. In general, we should use the optimal FKP weight $w = [1 + \bar{n}P]^{-1}$, which gives

$$L^{-3}_{\mathrm{eff}} = \frac{\int w^2\bar{n}^2\ \mathrm{d}^3r}{[\int w\,\bar{n}\ \mathrm{d}^3r]^2}. \tag{C.11}$$

Appendix D

Covariance matrices

We present here some example covariance matrices for the parameters p_i in the form of parameter standard deviations:

$$\sigma(p_i) = \sqrt{C_{ii}} \tag{D.1}$$

(no summation implied) and correlation matrices:

$$r_{ij} = C_{ij}/\sigma(p_i)\sigma(p_j) \tag{D.2}$$

for different combinations of values of β, slope n, and shot power to true power ratio at $k_{\rm max}$ $R = P_{\rm shot}/P_{\rm true}$. The standard deviations are scaled such that $\sigma(\beta) = 1$ and the number of modes N required to achieve this accuracy is given. For a fixed R the standard deviations scale as $N^{-1/2}$. For example, if the volume of the survey were increased by a factor of 100, and $k_{\rm max}$ and the number density of galaxies were kept constant, the errors on the parameters would decrease by a factor of 10. We show covariances for both the three-parameter model with $\sigma_{\rm p}'$ fixed and the full four-parameter model where it is allowed to vary. The number of modes is kept constant when the number of parameters is changed, so the value of $\sigma(\beta)$ reflects the additional uncertainty introduced. All models have $k_{\rm max}\sigma_{\rm p}' = 1$ and $F = 1$.

D.1 Example 1: $\beta = 1$; $n = -1.5$; $R = 0$; $N = 1120$

	β	F	$\ln A$
$\sigma(p)$	1.00	0.74	0.43
β	1.00	−0.76	−0.86
F	−0.76	1.00	0.56
$\ln A$	−0.86	0.56	1.00

Table D.1: Three-parameter model.

	β	F	σ'_{p}	$\ln A$
$\sigma(p)$	1.38	0.75	2.52	0.42
β	1.00	−0.43	0.69	−0.63
F	−0.43	1.00	0.16	0.55
σ'_{p}	0.69	0.16	1.00	−0.01
$\ln A$	−0.63	0.55	−0.01	1.00

Table D.2: Four-parameter model.

D.2 Example 2: $\beta = 0.3$; $n = -1.5$; $R = 0$; $N = 3210$

	β	F	$\ln A$
$\sigma(p)$	1.00	1.14	0.24
β	1.00	−0.96	−0.73
F	−0.73	1.00	0.58
$\ln A$	−0.73	0.58	1.00

Table D.3: Three-parameter model.

	β	F	σ'_{p}	$\ln A$
$\sigma(p)$	1.09	1.14	1.49	0.24
β	1.00	−0.87	0.41	−0.66
F	−0.87	1.00	−0.01	0.59
σ'_{p}	0.41	−0.01	1.00	0.00
$\ln A$	−0.66	0.59	0.00	1.00

Table D.4: Four-parameter model.

D.3 Example 3: $\beta = 1$; $n = -1$; $R = 0$; $N = 715$

	β	F	$\ln A$
$\sigma(p)$	1.00	1.01	0.53
β	1.00	−0.57	−0.86
F	−0.57	1.00	0.56
$\ln A$	−0.86	0.56	1.00

Table D.5: Three-parameter model.

	β	F	σ'_{p}	$\ln A$
$\sigma(p)$	1.59	1.03	3.13	0.53
β	1.00	−0.23	0.78	−0.55
F	−0.23	1.00	0.16	0.55
σ'_{p}	0.78	0.16	1.00	−0.01
$\ln A$	−0.55	0.55	−0.01	1.00

Table D.6: Four-parameter model.

D.4 Example 4: $\beta = 1$; $n = -1.5$; $R = 1000$; $N = 855\,000$

Note that, in this shot-noise dominated regime, the standard deviations scale with R to a good approximation for fixed N.

	β	F	$\ln A$
$\sigma(p)$	1.00	0.51	0.52
β	1.00	−0.85	−0.92
F	−0.85	1.00	0.65
$\ln A$	−0.92	0.65	1.00

Table D.7: Three-parameter model.

	β	F	σ'_{p}	$\ln A$
$\sigma(p)$	1.02	0.51	1.35	0.52
β	1.00	−0.82	0.17	−0.91
F	−0.82	1.00	0.03	0.65
σ'_{p}	0.17	0.03	1.00	0.00
$\ln A$	−0.91	0.65	0.00	1.00

Table D.8: Four-parameter model.

Appendix E

Window Functions

This appendix details the practical calculation of the various angular window matrices needed to construct the covariance matrix.

In this section, a compact notation is used to save repeatedly writing the indices (lm). Defining

$$Y' \equiv Y_{l'}^{m'} \tag{E.1}$$

and similarly for double primes etc. simplifies the notation considerably.

All integrations in this section are over 4π steradians of solid angle, so the infinitesimal $\mathrm{d}\Omega$ may be dropped from formulae for additional simplification.

E.1 Practical calculation of window functions

A huge amount of computer processor time is required to perform the repeated integrations to produce the full mask matrix, equation (4.40). However, a shortcut was discovered by Scharf *et al.* (1992). First, the spherical harmonic transform of the mask function is calculated,

$$M_{\ell m} = \int_{4\pi} M(\Omega) Y_{\ell}^{m*}(\Omega)\, \mathrm{d}\Omega. \tag{E.2}$$

This can be used to express the mask matrix in terms of integrals over three spherical harmonics,

$$W_{\ell\ell'}^{mm'} = \sum_{\ell''m''} M_{\ell''}^{m''} \int Y' Y'' Y^*. \tag{E.3}$$

Integrals over three spherical harmonics occur in the quantum mechanical theory of angular momentum. The integral can be expressed in terms of Clebsch-Gordon coefficients or, in this case, Wigner $3j$ matrices,

$$\int Y Y' Y'' = \sqrt{\frac{(2\ell+1)(2\ell'+1)(2\ell''+1)}{4\pi}} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ m & m' & m'' \end{pmatrix}.$$

For the integral to be non-zero, we must have $m+m'+m''=0$ and $|\ell-\ell'| \leq \ell'' \leq (\ell+\ell')$. Complex conjugates of spherical harmonics can be dealt with using the hermicity relation, $Y_\ell^{m*} = (-1)^m Y_\ell^{-m}$.

Fortunately, Wigner $3j$ coefficients can be expressed in terms of an algebraic formula; Landau & Lifschitz (1977) define them as

$$\begin{aligned} \begin{pmatrix} \ell & \ell' & \ell'' \\ m & m' & m'' \end{pmatrix} = \; & \delta_{m_1+m_2,m} \quad (-1)^{\ell-\ell'-m''} \sqrt{\frac{(\ell+\ell'-\ell'')!(\ell-\ell'+\ell'')!(-\ell+\ell'+\ell'')!}{(\ell+\ell'+\ell''+1)!}} \\ & [(\ell+m)!(\ell-m)!(\ell'+m')!(\ell'-m')!(\ell''+m'')!(\ell''+m'')!]^{1/2} \\ & \sum_z (-1)^z [z!(\ell+\ell'-\ell''-z)!(\ell-m-z)!(\ell'+m'-z)! \\ & (\ell''-\ell'+m+z)!(\ell''-\ell'-m'+z)!]^{-1}. \end{aligned} \tag{E.4}$$

The restrictions on the m and ℓ values prevent the factorials from becoming negative.

For terms with $m = m' = m'' = 0$,

$$\begin{aligned} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} &= (-1)^{L/2} \sqrt{\frac{(L-2\ell)!(L-2\ell')!(L-2\ell'')!}{(L+1)!}} \\ &\quad \times \frac{(L/2)!}{(L/2-\ell)!(L/2-\ell')!(L/2-\ell'')!}, \qquad \ell \quad \text{even}, \\ &= 0, \qquad \ell \quad \text{odd}. \end{aligned} \tag{E.5}$$

where $L = \ell + \ell' + \ell''$. These terms can also be calculated from a recurrence relation,

starting from $\begin{pmatrix} \ell & \ell & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$$\begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} = \sqrt{\frac{(L-2\ell'-1)(L-2\ell''+2)}{(L-2\ell')(L-2\ell''+1)}} \begin{pmatrix} \ell & \ell'+1 & \ell''-1 \\ 0 & 0 & 0 \end{pmatrix}. \quad \text{(E.6)}$$

Wigner $3j$ coefficients can also be expressed in terms of Clebsch-Gordon coefficients

$$\begin{pmatrix} \ell & \ell' & \ell'' \\ m & m' & m'' \end{pmatrix} = (-1)^{\ell-\ell'-m''}(2\ell''+1)^{-1/2}\langle \ell m \ell' m' | \ell \ell' \ell'' - m'' \rangle \quad \text{(E.7)}$$

Equivalent notations for these quantities exist in the literature, such as; $\langle \ell_1 m_1 \ell_2 m_2 | \ell_1 \ell_2 \ell m \rangle$, $\langle m_1 m_2 | \ell m \rangle$, $C(\ell_1, \ell_2, \ell; m_1, m_2)$ $C^{\ell}_{m_1 m_2}$ (Landau & Lifshitz 1977).

E.2 Real and imaginary parts

In the practical calculation of covariance matrices, the modes are divided into real and imaginary parts, see section (4.3.1). For the shot noise terms, it is necessary to create *Tesseral matrices*, formed from masked integrals over two Tesseral Harmonics (real and imaginary parts of spherical harmonics) *e.g.*

$$\Re e \Im m W^{mm'}_{ll'} = \int \Re e Y M \Im m Y', \quad \text{(E.8)}$$

with $\Re e \Re e \mathbf{W}$ etc. defined similarly. Note the ordering of the primed terms, $\Re e \Im m W \neq \Im m \Re e W$ in general, but instead the matrices are transposes of each other,

$$\Re e \Im m W^{mm'}_{ll'} = \Im m \Re e W^{m'm}_{l'l}, \quad \text{(E.9)}$$

so only three matrices need to be calculated explicitly: $\Re e \Re e \mathbf{W}$, $\Re e \Im m \mathbf{W}$ and $\Im m \Im m \mathbf{W}$.

The real and imaginary parts of the full $\mathbf{W}$ matrix can be expressed in terms of Tesseral harmonics,

$$\begin{aligned} \Re e \mathbf{W} &= \Re e \Re e \mathbf{W} + \Im m \Im m \mathbf{W} \\ \Im m \mathbf{W} &= \Re e \Im m \mathbf{W} - \Im m \Re e \mathbf{W}, \end{aligned} \quad \text{(E.10)}$$

so only three matrices need to be calculated from Wigner $3j$ coefficients. As with conventional **W** matrices, it would take a prohibitive amount of time to calculate these matrices explicitly. Fortunately it is possible to express Tesseral matrices in terms of triple integrals over spherical harmonics, allowing us to use Wigner $3j$ coefficients, as in section (E). Tesseral harmonics themselves can be expressed in terms of spherical harmonics, $\Re eY = (1/2)(Y + Y^*)$ and $\Im mY = (1/2i)(Y - Y^*)$, giving *e.g.*

$$\begin{aligned}
\Re e\Re eW_{ll'}^{mm'} &= \frac{1}{4}\int (Y+Y^*)M(\Omega)(Y'+Y'^*) \\
&= \frac{1}{4}(YMY' + YMY'^* + Y^*MY' + Y^*MY'^*) \\
&= \frac{1}{4}\sum_{l''m''} M_{l''}^{m''}\int (YY'Y'' + YY'^*Y'' + Y^*Y'Y'' + Y^*Y'^*Y'')
\end{aligned} \tag{E.11}$$

using equation (E.3). Similar algebra for the other Tesseral matrices yields

$$\Re e\Im mW_{ll'}^{mm'} = \frac{1}{4i}\sum_{l''m''} M_{l''}^{m''}\int (YY'Y'' - YY'^*Y'' + Y^*Y'Y'' - Y^*Y'^*Y'') \tag{E.12}$$

and

$$\Im m\Im mW_{ll'}^{mm'} = \frac{1}{4}\sum_{l''m''} M_{l''}^{m''}\int (-YY'Y'' + YY'^*Y'' + Y^*Y'Y'' - Y^*Y'^*Y''). \tag{E.13}$$

With the Condon-Shortley definition of spherical harmonics, *i.e.* $Y_\ell^{m*} = (-1)^m Y_\ell^{-m}$, the parity relations ($m \to -m$) for the Tesseral matrices are

$$\begin{aligned}
\Re e\Re eW_{m,-m'} &= (-1)^{m'}\Re e\Re eW_{mm'}, \\
\Re e\Re eW_{-m,m'} &= (-1)^{m}\Re e\Re eW_{mm'}, \\
\Re e\Re eW_{-m,-m'} &= (-1)^{m'+m}\Re e\Re eW_{mm'},
\end{aligned} \tag{E.14}$$

$$\begin{aligned}
\Im m\Im mW_{m,-m'} &= (-1)^{m'+1}\Im m\Im mW_{mm'}, \\
\Im m\Im mW_{-m,m'} &= (-1)^{m+1}\Im m\Im mW_{mm'}, \\
\Im m\Im mW_{-m,-m'} &= (-1)^{m+m'}\Im m\Im mW_{mm'},
\end{aligned} \tag{E.15}$$

$$\begin{aligned}
\Im m\Re eW_{m,-m'} &= (-1)^{m'}\Im m\Re eW_{mm'}, \\
\Im m\Re eW_{-m,m'} &= (-1)^{m+1}\Im m\Re eW_{mm'}, \\
\Im m\Re eW_{-m,-m'} &= (-1)^{m+m'+1}\Im m\Re eW_{mm'}.
\end{aligned} \tag{E.16}$$

Appendix F

Orthogonality Relations

This appendix lists the orthogonality relations for spherical harmonics and spherical Bessel functions, see Abramowitz & Stegun (1972) and Arfken (1985). With the boundary condition $j_l'(k_{ln}r_{\max}) = 0$, the orthogonality relation becomes (Abramowitz & Stegun 1972),

$$\int_0^{r_{\max}} j_l(k_{lm}r) j_l(k_{ln}r') r^2 \mathrm{d}r = \frac{\pi}{4k_{ln}^3}\delta^{\mathrm{K}}_{mn} \times \left[k_{ln}^2 r_{\max}^2 - \left(1+\frac{1}{2}\right)^2\right] J^2_{l+1/2}(k_{ln}r_{\max}) \tag{F.1}$$

while for any boundary condition the relation for Spherical Harmonics is

$$\int_{4\pi} Y_{lm}(\Omega) Y^*_{l'm'}(\Omega)\, d\Omega = \delta^K_{ll'}\delta^K_{mm'}, \tag{F.2}$$

with inverse

$$\sum_{lm} Y_{lm}(\Omega) Y^*_{lm}(\Omega') = \delta_D(\hat{\mathbf{x}} - \hat{\mathbf{x}}'). \tag{F.3}$$

REFERENCES

Abramowitz M., Stegun I.A., 1972, *Handbook of Mathematical Functions.* Dover, New York

Adler R.J., 1981, *The Geometry of Random Fields*, Wiley, Chichester

Alcock C., Paczyński B., 1979, Nature, 281, 358

Alcock C. *et al.* 1997, ApJ, 486, 697

Arfken G., 1985, *Mathematical Methods for Physicists.* Academic Press Inc., San Diego

Babul A., Weinberg D.H., Dekel A., Ostriker J.P., 1994, 427, 1

Ballinger W.E., Heavens A.F., Taylor A.N., 1995, MNRAS, 276, L59

Ballinger W.E., Peacock J.P., Hevens A.F., 1996, MNRAS, 282, 877

Bardeen J.M., Bond J.R., Kaiser N., Szalay A.S., ApJ, 304, 15

Baugh C.M., Efstathiou, G., 1993, MNRAS, 265, 145

Baugh C.M., Efstathiou, G., 1994, MNRAS, 267, 323

Bertschinger E., Dekel A., 1989, ApJ, 336, L5

Binney J., Quinn T., 1991, MNRAS, 249, 678

Binney J., Tremaine S., 1987, *Galactic Dynamics.* Princeton University Press, Princeton

Bond J.R., 1995, Phys. Rev. Lett., 74, 4369

Bond J.R., Efstathiou G., 1984, ApJ, 285 L45

Brandenberger R.H., 1990, in *Physics of the Early Universe*, proc 36th Scottich Universities Summer School in Physics, eds Peacock J.A., Heavens A.F., Davies A.T., Adam Hilger

Broadhurst T.J., Taylor A.N., Peacock J.A., 1995, ApJ, 438, 49

Bucher M., Goldhaber A.S., Turok N., 1995, Phys. Rev. D, 52, 3314

Bunn, E. F. 1995, Ph.D. Thesis, U.C. Berkeley, *ftp pac2.berkeley.edu/pub/bunn*

Bunn, E. F., Sugiyama N. 1995, ApJ, 446, 49

Carlberg R.G., Yee H.K.C., Ellington E., Abraham R., Gravel P., Morris S., Pritchet C.J., 1996, ApJ, 462, 32

Carr B.J., 1994, ARA&A, 32, 531

Carroll S.M., Press W.H., Turner L.E., 1992, ARA&A, 30, 499

Cole S., Fisher K.B., Weinberg D.H., 1994, MNRAS, 267, 785 (CFW94)

Cole S., Fisher K.B., Weinberg D.H., 1995, MNRAS, 275, 515 (CFW95)

Coles P., Lucchin F., *Cosmology: The Origin and Evolution of Cosmic Structure*, Wiley, Chichester

Copi C.J., Schramm D.N., Turner M.S., 1995, Science, 267, 192

Croft R.A.C., Efstathiou G., 1994a, MNRAS, 267, 390

Croft R.A.C., Efstathiou G., 1994b, MNRAS, 268, 23P

Davis M., Peebles P.J.E., 1983, ApJ, 267, 465

Davis M., Efstathiou G., Frenk C.S., White S.D.M., 1985, ApJS, 57, 241

Delel A., 1994, ARA&A, 32, 371

Dekel A., 1997, preprint astro-ph/9705033

Dekel A., Rees M.J., 1987, Nat, 326, 455

Dekel A., Bertschinger E., Faber S.M., 1990, ApJ, 364, 349

Dekel A., Burstein D., White S.D.M., 1997, in *Critical Dialogues in Cosmology*, ed. Turok N., World Scientific Press, Singapore

de Laix A.A., Starkman G., 1997, preprint astro-ph/9707008

Dunlop J.S., Peacock J.A., Jimenez R., Dey A., Spinrad H., Stevens D., Windhorst R., 1996, Nat.,381, 581

Efstathiou G.P., 1990, in *Physics of the Early Universe*, proc 36$^{\text{th}}$ Scottich Universities Summer School in Physics, eds Peacock J.A., Heavens A.F., Davies A.T., Adam Hilger

Efstathiou G.P., 1995, MNRAS, 274, 73P

Efstathiou G.P., 1995, MNRAS, 276, 1425

Efstathiou G.P., 1996, in *Cosmology and Large Scale Structure*, Les Houches Session LX, eds Shaeffer R., Silk J., Spiro M., Jinn Justin J., Elsevier Science Publishers, The Netherlands

Efstathiou G.P., Ellis R.S., Peterson B.A., 1988, MNRAS, 232, 431

Efstathiou G.P., Sutherland W.J., Maddox S.J., 1990, Nat, 348, 705

Efstathiou G.P., Davis M., Frenk C.S., White S.D.M., 1985, ApJS, 57, 241

Fabbri R., Natale V., 1990, ApJ, 363, 3

Faber S.M., Gallagher J.S., 1979, ARA&A, 17, 135

Fan Z., Bardeen J.M., 1995, Phys. Rev. D, 51, 6714

Feldman H.A., Kaiser N., Peacock J.A., 1994, ApJ, 426, 23 (FKP)

Fisher K.B., Nusser A., MNRAS, 1996, 279, L1

Fisher K.B., Davis M., Strauss M.A., Yahil A., Huchra J.P., 1993, ApJ, 402, 42

Fisher K.B., Scharf C.A., Lahav O., 1994, MNRAS, 266, 219

Fisher K. B., Huchra J. P., Strauss M. A., Davis M., Yahil A., Schlegel D., 1995a, ApJS, 100, 69

Fisher K.B., Lahav O., Hoffman Y., Lynden-Bell D., Zaroubi S., 1995, MNRAS, 272, 885

Gamow, G., 1946, Phys. Rev., 70, 527.

Gunn J., Weinberg, D., 1995, in Maddox S.J., Aragón-Salamanca A., eds, Wide-field Spectroscopy and the Distant Universe, Proc. 35th Herstmonceux Conference. World Scientific, Singapore, p.3

Hamilton A.J.S., 1992, ApJ, 385, L5

Hamilton A.J.S., 1993, ApJ, 406, L47

Hamilton A.J.S., 1995, in *Clustering in the Universe*, Proc. XVth Rcontres de Moriond, ed. Maurogordato S., Balkowski C., Tao C., Trân Thanh Vân J., Editions Frontières, pp 143-155

Hamilton A.J.S., 1997, preprint astro-ph/9708102

Hatton S.J., Cole S., 1997, MNRAS submitted, preprint astro-ph/9707186

Hamilton A.J.S., Culhane M., 1996, MNRAS, 278, 73

Hawkins M.R.S., 1993, Nat, 366, 242

Heavens A.F., Taylor A.N., 1995, MNRAS, 275, 483 (HT)

Hubble E., 1929, Proc. Nat. Acad., 15, 168

Jacoby G.H. *et al.* 1992, PSAP, 104, 599

Jackson J.C., 1972, MNRAS 156, 1P

Jackson J.D., 1975, Classical Electrodynamics. John Wiley & Sons, New York

Jing Y.P., Mo H.J., Börner G., 1997, preprint astro-ph/9707106

Jones M.E., Saunders R., Baker J.C., Cotter G., Edge A., Grainge K., Haynes T., Lasenby A., Pooley G., Röttgering H., 1997, ApJ, 479, L1

Jungman, G., Kamionkowski, M., Kosowsky A. & Spergel D. N. 1996, Phys. Rev. D, 54, 1332

Kaiser N., 1987, MNRAS, 227, 1

Kaiser N., Squires G., 1993, ApJ, 404, 441

Karhunen, K., *Über lineare Methoden in der Wahrscheinlichkeitsrechnung* (Kirjapaino oy. sana, Helsinki, 1947).

Kendall M. G., Stuart, A., 1969. The Advanced Theory of Statistics, Volume II, Griffin, London

Kenney, J. F. & Keeping, E. S. 1951, *Mathematics of Statistics, Part II*, 2nd ed. (Van Nostrand, New York).

Klypin A., Holtzman J., Primack J., Regös E., 1993, ApJ, 416, 1

Kolb E.W., Turner M.S., 1990, *The Early Universe*, Adisson-Wesley, USA

Lahav O., 1993, in Bouchet F.R., Lachiéze-Rey M., eds, Proc. 9th IAP Conference, Cosmic Velocity Fields. Editions Frontières, Gif-Sur-Yvette, p.205

Landau L.D., Lifshitz E.M., 1977, Quantum Mechanics, 3rd ed., Permagon Press, Oxford

Liddle A.R., Lyth D.H., 1992,

Lilje P., Efstathiou G.P., 1989, MNRAS, 236, 851

Linde A.D., 1983, Phys. Lett., 129B, 177

Longair M.S., 1984, *Theoretical Concepts in Physics*, Cambridge University Press, Cambridge

Lynden-Bell D., Lahav O., Burstein D., 1989, MNRAS, 241, 325

Madau P., Ferguson H.C., Dickonson M.E., Giavilsco M., Steidel C.C., Fruchter A., 1996, MNRAS, 283, 1388

Mann R.G., Peacock J.A., Heavens A.F., 1997, preprint astro-ph/9708020

Mann R.G., Saunders W., Taylor A.N., 1996, 279, 636

Matarrese S., Verde L., Heavens A.F., 1997, preprint astro-ph/9706059

Mather J.C. *et al.* 1994, ApJ, 420, 439

Matsubara T., Suto Y., 1996, ApJ, 470, 1

Meiksin A., Davis M., 1986, Astr. J., 91, 191

Mo H.J., Jing Y.P., Börner G., 1993, MNRAS, 264, 825

Nusser A., Davis M., 1994, MNRAS, 421, L1

Pagel B.E., Simonson E.A., Terlevich R.J., Edmunds M.G., 1992, MNRAS, 255, 325

Peacock J.A., 1992, in Martinez V., Portilla M., Sáez D., eds, New insights into the Universe, Proc. Valencia Summer School. Springer, Berlin, p.1

Peacock J.A., 1992, MNRAS, 258, 581

Peacock J.A., 1997, MNRAS, 284, 885

Peacock J.A., Nicholson D., 1991, MNRAS, 253, 307

Peacock J.A., Dodds S.J., 1994, MNRAS, 267, 1020

Peebles P.J.E., 1973, ApJ, 185, 413

Peebles P.J.E., 1976, Ap Space Sci., 45, 3

Peebles P.J.E., Groth E.J., 1976, A&A, 53, 131

Peebles P.J.E., 1980, *The Large Scale Structure of the Universe.* Princeton University Press, Princeton

Peebles P.J.E., 1993, *The Principles of Physical Cosmology,* Princeton University Press, Princeton

Penzias A.A., Wilson R.W., 1965, ApJ, 142, 419

Perlmutter *et al.* 1997, ApJ, 483, 565

Phillipps S., 1994, MNRAS, 269, 1077

Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P., 1992, Numerical Recipes. Cambridge University Press, Cambridge

Rowan-Robinson M. *et al.* 1990, MNRAS, 274, 1

Ryden B., 1995, ApJ, 452, 25

Sachs R.K., Wolfe A.M., 1967, ApJ, 147, 73

Sargent W.L.W., Turner E.L., 1977, ApJ, 212, L3

Saunders W., Rowan-Robinson M., Lawrence A., 1992, MNRAS, 258, 134

Saunders W. *et al.* 1995, in Maddox S.J., Aragón-Salamanca A., eds, Wide-field Spectroscopy and the Distant Universe, Proc. 35th Herstmonceux Conference. World Scientific, Singapore, p.88

Scharf C.A., Lahav O., 1993, MNRAS, 264, 439

Scharf C.A., Hoffman Y., Lahav O., Lynden-Bell D., 1992, MNRAS, 259, 229

Shanks T., Boyle B.J., 1994, MNRAS, 271, 753

Sheth R.K., 1996, MNRAS, 279, 1310

Sigad Y, Amiram E., Dekel A., Strauss M.A., Yahil A., 1997, preprint astro-ph/9708141

Silk J., 1967, Nat, 215, 1155

Smoot G.F. *et al.* 1992, ApJ, 396, L1

Stirling A.J., Peacock J.A., 1996, MNRAS, 283, 99P

Strauss M.A., Willick J.A., 1995, Phys. Rep., 261, 271

Strauss M.A., Davis M., Yahil A., Huchra J.P., 1992, ApJ, 385, 421

Tadros H., Efstathiou G.P., 1995, MNRAS, 276, 45P

Tadros H., Ballinger W.E., Heavens A.F., Taylor A.N., Efstathiou G., Frenk C.S., Keeble O., MacMahon R., Maddox S.J., Oliver S., Rowan-Robinson M., Saunders W., Sutherland W.J., White S.D.M., in preparation, for submission to MNRAS

Taylor A.N. & Hamilton A.J.S., 1996, MNRAS, 282, 767

Taylor A.N., Rowan-Robinson M., 1992, Nat, 359, 396

Taylor K., 1995, in Maddox S.J., Aragón-Salamanca A., eds, Wide-field Spectroscopy and the Distant Universe, Proc. 35th Herstmonceux Conference. World Scientific, Singapore, p.15

Tegmark, M. 1995, ApJ, 455, 429

Tegmark M., Taylor A.N., Heavens A.F., 1997, ApJ, 480, 22

Trimble V., 1987, ARA&A, 25, 425

Tytler D., Fan X.M., Burles S., Nat, 381, 207

van Dalen A., Schaefer R.K., 1992, ApJ, 398, 33

Vilenkin A., Shellard E.P.S., 1994, *Cosmic Strings and other Topological Defects*, Cambridge University Press, Cambridge

Vogeley, M. S. 1995, in "Wide-Field Spectroscopy and the Distant Universe", eds. Maddox & Aragón-Salamanca (World Scientific, Singapore)

Vogeley, M. S. & Szalay, A. S., 1996, preprint astro-ph/9601185

Wagoner, R.V., Fowler, W.A., & Hoyle F., 1967, ApJ, 148, 3

Walker T.P., Steigman G., Scramm D.N., Olive K.A., Kang H.S., 1991, ApJ, 376, 51

Weinberg S., 1972, Gravitation & Cosmology. Wiley, New York

Weinberg S., 1989, Rev. Mod. Phys., 61, 1

White S.D.M., Frenk C.S., Davis M., 1983, ApJ, 271, L1

White S.D.M., Davis M., Frenk C.S., 1984, MNRAS, 209, 27P

White S.D.M., Scott D., Silk J., 1994, ARA&A, 32, 319

Willick J.A., Strauss M.A., Dekel A., Kolatt T., 1997, ApJ, 486, 629

Wright E.L., Bennet C.L., Gorski K., Hinshaw G., Smoot G.F., 1996, ApJ, 464, L21

Yahil A., Walker D., Rowan-Robinson M., ApJ, 301, L1

Zaroubi S., Hoffman Y., 1996, ApJ, 462, 25

Zeldovich, Ya.B., 1970, A&A, 5, 84.

Zwicky, F., 1933, Helv. Phys. Acta, 6, 110.